An Examination of the Static and Dynamic Performance of Interest Rate Option Pricing Models In the Dollar Cap-Floor Markets

Anurag Gupta[a*]

Marti G. Subrahmanyam[b*]

Current version: September 2001

JEL Classification: G12; G13; G19
Keywords: Interest rate options; interest rate caps / floors; term structure of interest rates

[a]Department of Banking and Finance, Weatherhead School of Management, Case Western Reserve University, 10900 Euclid Avenue, Cleveland, Ohio 44106-7235. Ph: (216) 368-2938, Fax: (216) 368-4776, E-mail: axg77@po.cwru.edu.
[b]Department of Finance, Leonard N. Stern School of Business, New York University, New York, NY 10012-1126. Ph: (212) 998-0348, Fax: (212) 995-4233, E-mail: msubrahm@stern.nyu.edu.
[*] We thank Viral Acharya, Andrew Carverhill, Stephen Figlewski, Kenneth Garbade, A. R. Radhakrishnan, Matthew Richardson, Peter Ritchken, and Dick Stapleton for valuable comments on earlier drafts. We are also thankful to the participants in the seminars at Baruch College, Case Western Reserve University, McGill University, New York University, Rutgers University, University of South Carolina, University of Strathclyde, University of Toronto, and conference participants at the Derivatives Securities Conference at Boston University, the Western Finance Association meetings, the European Financial Management Association meetings, the European Finance Association meetings and the Financial Management Association meetings for comments and suggestions. The usual disclaimer applies.

An Examination of the Static and Dynamic Performance of Interest Rate Option Pricing Models In the Dollar Cap-Floor Markets

Abstract

This paper examines the static and dynamic accuracy of interest rate option pricing models in the U.S. dollar interest rate cap and floor markets. We evaluate alternative one-factor and two-factor term structure models of the spot and the forward interest rates on the basis of their out-of-sample predictive ability in terms of pricing and hedging performance. The one-factor models analyzed consist of two spot-rate specifications (Hull and White (1990) and Black-Karasinski (1991), five forward rate specifications (within the general Heath, Jarrow and Morton (1990b) class), and one LIBOR market model (Brace, Gatarek and Musiela (1997) [BGM]). For two-factor models, two alternative forward rate specifications are implemented within the HJM framework. We conduct tests on daily data from March-December 1998, consisting of actual cap and floor prices across both strike rates and maturities. Results show that fitting the skew of the underlying interest rate distribution provides accurate *pricing* results within a one-factor framework. However, for *hedging performance*, introducing a second stochastic factor is more important than fitting the skew of the underlying distribution. Overall, the one-factor lognormal model for short term interest rates outperforms other competing models in pricing tests, while two-factor models perform significantly better than one-factor models in hedging tests. Modeling the second factor allows a better representation of the dynamic evolution of the term structure by incorporating expected twists in the yield curve. Thus, the interest rate dynamics embedded in two-factor models appears to be closer to the one driving the actual economic environment, leading to more accurate hedges. This constitutes evidence against claims in the literature that correctly specified and calibrated one-factor models could replace multi-factor models for consistent pricing and hedging of interest rate contingent claims.

1. Introduction

Interest rate option markets are amongst the largest and most liquid option markets in the world today, with daily volumes of billions of U.S. dollars in trading of interest rate caps/floors, Eurodollar futures options, Treasury bond futures options, and swaptions. The total notional principal amount of over-the-counter interest rate options such as caps/floors and swaptions outstanding at the end of 2000 was about $9.5 trillion.[1] These options are widely used both for hedging as well as speculation against changes in interest rates.

Theoretical work in the area of interest rate derivatives has produced a variety of models and techniques to value these options, some of which are widely used by practitioners.[2] The development of many of these models was mainly motivated by their analytical tractability. Therefore, while these models have provided important theoretical insights, their empirical validity and performance remain to be tested. Empirical research in this area has lagged behind theoretical advances partly due to the difficulty in obtaining data, as most of these interest rate contingent claims are traded in over-the-counter markets, where data are often not recorded in a systematic fashion. This gap is being slowly filled by recent research in this area.

This paper provides empirical evidence on the validity of alternative interest rate models. We examine the static and dynamic accuracy of interest rate option pricing models in the U.S. dollar interest rate cap and floor markets. For the first time in this literature, a time series of actual cap and floor prices across strike rates and maturities is used to study the systematic patterns in the pricing and hedging performance of competing models, on a daily basis. Alternative one- and two-factor models of the term structure are evaluated based on their static performance (by examining their out-of-sample price predictions) and their dynamic accuracy (by analyzing their ability to hedge caps and floors). The one-factor models analyzed consist of two spot-rate specifications (Hull and White (1990) [HW] and Black-Karasinski (1991) [BK]), five forward rate specifications (within the general Heath, Jarrow and Morton (1990b) [HJM] class), and one LIBOR market model (Brace, Gatarek and Musiela (1997) [BGM]). For two-factor models, two alternative forward rate

[1] Source: Bank for International Settlements (BIS) Quarterly Review, December 2000.

[2] The early models, many of which are still widely used, include those by Black (1976), Vasicek (1977), Cox, Ingersoll and Ross (1985), Ho and Lee (1986), Heath, Jarrow and Morton (1990b), Hull and White (1990), Black, Derman and Toy (1990), and Black and Karasinski (1991). Several variations and extensions of these models have been proposed in the literature in the past decade.

specifications are implemented within the HJM framework. The analysis in this paper, therefore, sheds light on the empirical validity of a broad range of models for pricing and hedging interest rate caps and floors, especially across different strikes, and suggests directions for future research.

There are very few papers that study the empirical performance of these models in valuing interest rate derivatives. Flesaker (1993) and Amin and Morton (1994) test the HJM model in pricing Eurodollar future options. The Amin and Morton (1994) study evaluates different volatility specifications within the HJM framework, using a time-series of Eurodollar futures and options data. They document systematic strike rate and time-to-maturity biases for all models. However, their analysis is restricted to options with relatively short maturities (less than one-year), in a market with much lower trading volumes than those for caps / floors. Therefore, their analysis does not capture the longer-term effects of the volatility term structure, including mean-reversion. Also, they do not evaluate any spot rate specifications, and restrict their analysis to single factor models. Canabarro (1995) examines the accuracy of interest rate hedges constructed using the Black-Derman-Toy and two-factor extensions of the Cox-Ingersoll-Ross and Brennan-Schwartz models, and finds that two-factor bond replicating strategies are more accurate than one-factor ones. However, his study is based on simulated data on Treasury yield curves, and does not examine many of the more recent term structure models. Bühler, Uhrig, Walter and Weber (1999) test different one-factor and two-factor models in the German fixed-income warrants market. In their comprehensive study, they report that the one-factor forward rate model with linear proportional volatility outperforms all other models. Their study, based on weekly data, is limited to options with maturities of less than 3 years. In addition, the underlying asset for these options is not homogenous. For some of the options, the underlying asset is the ten-year German Treasury bond ("the BUND"), while for others, it is the five-year German Treasury bond ("the BOBL"). The methodology in this study involves the estimation of model parameters from historical interest rate data rather than the extraction of this information from derivative prices. Therefore, the results are subject to large pricing errors. Lastly, the paper does not analyze strike-rate biases, due to data limitations. However, casual observation and evidence from other derivative markets suggest that these biases may be significant.

There have been some recent working papers that test model performance for pricing interest rate derivatives. Ritchken and Chuang (1999) test a three-state Markovian model in the Heath-Jarrow-Morton paradigm when the volatility structure of forward rates is humped, using price data for at-the-money (ATM) caplets. They find that with three state variables, the model captures the full dynamics of the term structure without using any time varying parameters. However, a single state variable model is unable to achieve such a fit. They conclude that the volatility hump is an important feature to be captured in a term structure model. Hull and White (1999) test the LIBOR market model for swaptions and caps across a range of strike rates, but with data for only one day, August 12, 1999. They find that the absolute percentage pricing error for caps was greater than for swaptions. Longstaff, Santa-Clara and Schwartz (2001, LSS) use a string model framework to test the relative valuation of caps and swaptions using ATM cap and swaptions data. Their results indicate that swaption prices are generated by a four-factor model, and that cap prices periodically deviate from the no-arbitrage values implied by the swaption market. Moraleda and Pelsser (2000) test three alternative spot-rate models and two Markovian forward-rate models on cap and floor data from 1993-94, and find that spot rate models outperform the forward-rate models. However, as they acknowledge, their empirical tests are not very formal.

None of the above-mentioned papers examines the hedging performance of the alternative models, except the one by LSS where they test their four factor model against the Black model, and show that the performance of the two models is statistically indistinguishable, and a recent paper by Driessen, Klassen and Melenberg (2000, DKM) whose analysis runs parallel to the direction of our paper. DKM test one-factor and multi-factor HJM models with respect to their pricing and hedging performance using ATM cap and swaption volatilities. They find that a one-factor model produces satisfactory pricing results for caps and swaptions. In terms of hedging performance, for both caps and swaptions, they find that the choice of hedge instruments affects the hedging accuracy more than the particular term structure model chosen. However, as with all other studies cited above, their data set is restricted to ATM options. As noted earlier, the strike rate effect may be important since many of the model imperfections are more evident when one analyzes options away-from-the-money. While it is interesting that they find satisfactory pricing and hedging performance using a one-factor model, even for swaptions, their results are not surprising. The question is whether this conclusion holds up for options that are away-from-the-money. In our paper, we specifically focus on cap and floor prices across *different* strike rates and maturities, to examine how alternative term structure models are affected by strike biases.[3]

---

[3] Another recent paper by Andersen (1999) adapts a multi-factor LIBOR market model to price Bermudan swaptions using simulations; however, it is not related to the empirical issues that we address in this

In this paper, the empirical performance of analytical models is evaluated along two dimensions – their static and dynamic accuracy. Static performance refers to their ability of a model to price options accurately at a given point in time, given that the model is estimated in a manner that is consistent with market observables. Dynamic accuracy refers to the ability of the model to capture movements in the term structure after being initially calibrated to fit market observables. The static accuracy of a model is useful in picking out deviations from arbitrage-free pricing. As for dynamic accuracy, the correct representation of the behavior of the term structure of interest rates is a crucial feature to validate an arbitrage-free model as an accurate tool to hedge interest rate claims. The hedging tests examine whether the interest rate dynamics embedded in the model is similar to that driving the actual economic environment that the model is intended to represent.

Our results show that, for plain-vanilla interest rate caps and floors, a one-factor lognormal forward rate model outperforms other competing one-factor models, in terms of out-of-sample pricing accuracy. In addition, the estimated parameters of this model are stable. In particular, the one-factor BGM model outperforms other models in pricing tests where the models are calibrated using option pricing data for the same day for which they are used to estimate prices of other options. We also find that the assumption of lognormally distributed interest rates results in a smaller "skew" in pricing errors across strike rates, as compared to other distributions assumed in alternative interest rate models. Two-factor models improve pricing accuracy only marginally. Thus, for accurate pricing of caps and floors, especially away-from-the-money, it is more important for the term structure model to fit the skew in the underlying interest rate distribution, than to have a second stochastic factor driving the term structure. However, the hedging performance improves significantly with the introduction of a second stochastic factor in term structure models, while fitting of the skew in the distribution improves hedging performance only marginally. This occurs because two-factor models allow a better representation of the dynamic evolution of the yield curve, which is more important for hedging performance, as compared to pricing accuracy. Thus, even for simple interest rate options such as caps and floors, there is a significant advantage to using two-factor models, over and above fitting the skew in the underlying (risk-neutral) interest rate distribution, for consistent pricing and hedging within a book. This refutes claims in the literature that correctly specified and calibrated one-factor models could eliminate the need to have multi-factor models for pricing and hedging interest rate derivatives.[4]

---

paper.

[4] For instance, Hull and White (1995) state that "the most significant difference between models is a strike

We examine two alternative calibrations of the spot rate models. In the first implementation, the volatility and mean-reversion parameters are held constant. As a result, while the models are calibrated to fit the current term structure exactly, the model prices match the current cap/floor prices only with an error, albeit by minimizing its impact. In the alternative implementation, an additional element of flexibility is introduced by making the parameters time-varying. This enables us to fit both the current term structure and the cap/floor prices exactly, although this renders the parameter estimates unstable. Thus, there is a tradeoff between the imperfect fit of the models and the instability of the model parameters, which is examined in our empirical analysis.

The paper is organized as follows. Section 2 presents an overview of the different term structure models used for pricing and hedging interest rate contracts. In section 3, details of estimation and implementation of these term structure models are discussed. Section 4 describes the design of this empirical study and the different methodologies used in evaluating the alternative models. Section 5 describes the data used in this study, along with the method used for constructing the yield curve. The results of the study are reported in section 6. Section 7 concludes.

2. Overview of term structure models for pricing caps/floors

The interest-rate derivatives market consists of instruments that are based on different market interest rates. Interest rate swaps and FRAs are priced based on the *level* of different segments of the yield curve; caps and floors are priced based on the level and the *volatility* of the different forward rates (i.e., the diagonal elements of the covariance matrix). Swaptions are priced based on both the diagonal and the off-diagonal elements of the same covariance matrix, i.e., they also price the correlations among the forward rates. Since caps and floors do not price the correlations among forward rates, it appears, at first glance, that one-factor models might be sufficiently accurate in pricing them, and the added numerical complexity of multi-factor models (in particular, two-factor models) may not be justified.[5] This is also one of the key issues that this paper seeks to investigate.

price bias ... the number of factors in a term structure model does not seem to be important except when pricing spread options ... one-factor Markov models when used properly do a good job of pricing and hedging interest-rate sensitive securities".

[5] One-factor term structure models imply perfectly correlated spot/forward rates, while two-factor (and multi-factor) models allow for imperfect correlation between spot/forward rates of different maturities.

There are a large number of term structure models for the valuation of interest-rate derivatives. They can broadly be categorized into two groups. The first one models the dynamics of the instantaneous or discrete-time spot interest rate (spot rate models), and the second models the arbitrage-free evolution of the entire term structure of forward rates (forward rate models).

In the first group of models (spot rate models), the entire term structure is inferred from the evolution of the spot short-term interest rate (and, in case of two-factor models, by another factor such as the long-term interest rate, the spread, the volatility factor, or the futures premium). This includes the traditional models by Vasicek (1977), Brennan and Schwartz (1979), Cox, Ingersoll and Ross (1985), Longstaff and Schwartz (1992), Stapleton and Subrahmanyam (1999) and others. However, the equilibrium models such as those by Vasicek (1977), Brennan and Schwartz (1979) and Cox, Ingersoll and Ross (1985) determine the term structure endogenously; hence, they do not fit the current term structure exactly. This implies that the models may permit arbitrage opportunities across zero coupon bonds, even prior to pricing derivatives. Given the resulting potential mispricing of the underlying discount bonds, the error introduced in the prices of derivatives based on these bonds may be accentuated, because of their inability to price derivatives satisfactorily. These models could be modified to match the term structure *exactly* in an arbitrage-free framework by making one or more of the parameters time-varying. This is implemented in the models by Hull and White (1990), Black, Derman and Toy (1990), Black and Karasinski (1991), Peterson, Stapleton and Subrahmanyam (1999) and others. These no-arbitrage models take the current term structure as an input rather than an output, thus making the yield curve consistent with the observed prices of discount bonds.

The approach of modeling the forward, rather than the spot, interest rates was pioneered by Ho and Lee (1986). Ho and Lee take as given the prices of discount bonds of all maturities and model the subsequent evolution of this price vector to preclude arbitrage opportunities. This is equivalent to modeling the forward interest rate curve, which was the approach used by HJM (1990b) in extending and generalizing the work of Ho and Lee in a continuous time framework. HJM model the instantaneous forward-rate curve with a fixed number of unspecified factors that drive the dynamics of these forward-rates. The form of the forward rate changes can be specified in a fairly general manner. In fact, many of the processes specified for the evolution of the spot interest rate can be treated as special cases of HJM models by appropriately specifying the volatility function of the forward interest rates. For example, specifying the volatility as an exponential function of the time to maturity gives rise to the Ornstein-Uhlenbeck process as in Vasicek (1977). A constant volatility results in the continuous time version of the Ho and Lee model. In these two cases, closed form solutions are available for discount bonds and option prices.

In recent years, the so-called “market models” have become very popular amongst practitioners. These models recover market pricing formulae by the direct modeling of market quoted rates. This approach overcomes one of the drawbacks of the traditional HJM models: that they involve instantaneous forward rates that are not directly observable (and are hence difficult to calibrate). A model that is popular among practitioners is the one proposed by Brace, Gatarek and Musiela (1997) [BGM].[6] They derive the processes followed by market quoted rates within the HJM framework, and deduce the restrictions necessary to ensure that the distribution of market quoted rates of a *given* tenor under the risk-neutral forward measure is log-normal. With these restrictions, caplets of that tenor satisfy the Black (1976) formula for options on forward / futures contracts.

In spot rate models, all the rates are derived from the evolution of the spot rate. In order to incorporate realistic correlation levels across the term structure, additional factors have then to be introduced in the form of another stochastic variable such as the long term rate, short rate volatility, the slope of the term structure, the mean-reversion parameter, etc. In contrast, the HJM framework allows the forward rates maturing at various fixed points in time to evolve simultaneously. The forward rate curve evolution can be modeled as being driven by any number of stochastic variables or factors.[7] In theory, each of the forward rates could be driven by a separate stochastic variable yielding as many factors as there are forward rates. This allows the incorporation of correlations through appropriate specification of the volatility functions for each of the factors.

In this paper, we analyze the comparative performance of various one-factor and two-factor spot rate, forward rate and market models. The spot rate models analyzed are the one-factor HW and BK models. In the forward rate class, one-factor and two-factor models are considered. The HJM framework is used to implement different assumptions about the distribution of the underlying forward rate, through appropriate specification of the volatility functions.[8] Amongst the market models, the one-factor BGM model is analyzed.

[6] A similar model has also been proposed by Miltersen, Sandmann and Sondermann (1997).
[7] In theory, one could also model the spot rates with a multi-factor specification.
[8] In the HJM framework, the two-factor model nests the corresponding one-factor model, thus making it easier to compare the results of the two alternative specifications and infer the impact of introducing a

2.1 Spot rate models

There is a large variety of spot rate models in the literature. They can be adapted to the current term structure of interest rates and volatilities by making the parameters of the stochastic processes time-dependent. These time-dependent parameters are determined in a way such that both the endogenous term and volatility structures fit the observed ones exactly.

A generalized one-factor spot rate specification, that explicitly includes mean reversion, has the form:

$$df(r) = [q(t) - af(r)]dt + s\,dz \qquad (1)$$

where

f(r) = some function *f* of the short rate *r*,

q(t) = a function of time chosen so that the model provides an exact fit to the initial term structure, usually interpreted as a time-varying mean,

a = mean- reversion parameter,

s = volatility parameter.

Two special cases of the above model are in widespread use. When *f(r)=r*, the resultant model is the HW model (also referred to as the extended-Vasicek model)

$$dr = [q(t) - ar]dt + s\,dz \qquad (2)$$

f(r)=ln(r) leads to the BK model

$$d\ln r = [q(t) - a\ln r]dt + s\,dz \qquad (3)$$

The volatility parameter, *s*, determines the overall level of volatility, while the reversion parameter, *a*, determines the relative volatilities of long and short rates. The probability distribution of short rate is Gaussian in the HW model and lognormal in the BK model.

In this paper, these models are estimated in two different ways. In the first implementation, the mean-reversion parameter '*a*' and the short rate volatility '*s*' are both held constant. Therefore, the models are estimated with only one time-dependent parameter such that it fits the current term structure exactly. The remaining parameters of the process are determined so as to achieve a 'best

---

second stochastic factor.

fit' to the observed volatility term structure. The drawback with this implementation is that by keeping the reversion and volatility parameters constant, the model does not fit the current cap / floor prices exactly, which induces an inherent mispricing to start with. The advantage of keeping the parameters constant is the resulting stability of parameter estimates as well as the stationarity of the volatility term structure.

To understand the effect of making the parameters time varying, the second implementation of these models is conducted by making the reversion and volatility parameters time varying. This allows the models more degrees of freedom to make the current interest rate tree fit the prices of caps / floors as well. However, fitting to option prices has implications for the future evolution of the term structure. Making two or more parameters time varying may result in unstable parameter estimates and implausible future evolutions of the term structure.[9] This would be reflected in poor out-of-sample performance of these models. Hence, there is a tradeoff between a perfect fit of the current term structure of volatility and the stationarity of the model parameters.[10]

2.2 Forward rate models

In the forward rate models, the HJM framework allows the valuation of contingent claims without having to estimate the market price of risk or any drift parameters. The drift is completely defined by the volatility parameters. By appropriately specifying the volatility structure, virtually any interest rate distribution can be studied. This framework lends itself very well to the comparative evaluation of one-factor and two-factor models as the two-factor model nests the one-factor model, which can be easily obtained by setting the second volatility parameter to zero. Hence, a single estimation of the volatility parameters is sufficient to implement both the models. All other models require separate estimation of the model parameters for the one-factor and two-factor versions. Also, the HJM framework matches the current term structure, by construction; hence, it does not lead to mispricing the underlying discount bonds.

Let $f(t,T)$ be the forward interest rate at date t for instantaneous riskless borrowing or lending at date T. The HJM approach models the evolution of the entire instantaneous forward rate curve, driven

[9] This non-stationarity would be more problematic for instruments whose prices depend on future volatility term structures (like American / Bermudan options, spread options captions, etc.). For standard caps and floors, as in this paper, this is less important.

by a fixed number of unspecified factors. Forward interest rates of every maturity T evolve simultaneously according to the stochastic differential equation

$$df(t,T) = m(t,T,.)dt + \sum_{i=1}^{n} s_i(t,T,\ f(t,T))dW_i(t) \tag{4}$$

Where $W_i(t)$ are n independent one-dimensional Brownian motions and $m(t,T,.)$ and $s_i(t,T,f(t,T))$ are the drift and volatility coefficients for the forward interest rate of maturity T. [11] The volatility coefficient represents the instantaneous standard deviation (at date t) of the forward interest rate of maturity T, and can be chosen arbitrarily. For each choice of volatility functions $s_i(t,T,f(t,T))$, the drift of the forward rates under the risk-neutral measure is uniquely determined by the no-arbitrage condition

$$m(t,T,.) = \sum_{i=1}^{n} s_i(t,T,\ f(t,T))\int_t^T s_i(t,s,\ f(t,s))ds \tag{5}$$

The drift term for the forward rate maturing at T depends on the instantaneous standard deviation of all forward rates maturing between t and T. The choice of the volatility function $s_i(t,T,f(t,T))$ determines the interest rate process that describes the stochastic evolution of the entire term structure. If the volatility function is stochastic, it *may* make the interest rate process non-Markovian, in which case no closed-form solutions are possible for discount bonds or options.[12] Hence, it is necessary to restrict the nature of the volatility functions in order to obtain manageable solutions.

The volatility functions analyzed in this paper, $s_i(t,T,f(t,T))$, are time invariant functions. In these functions, the volatility depends on t and T only though T-t. Therefore, given a term structure at time t, the form of its subsequent evolution through time depends only on the term structure, not on the specific calendar date t. Even with this restriction, a rich class of volatility structures can be analyzed. To preserve the stability of parameter estimation, we analyze only one- and two-parameter volatility functions in this paper. Hence, we focus on the following volatility functions, and models that they imply:

---

[10] See Hull and White (1996) for a discussion on this issue.

[11] The drift coefficient for each maturity T depends on forward interest rates of all other maturities, the dependence being represented by "." as the third argument of $m(t,T,.)$.

[12] As discussed later, Ritchken and Sankarasubramanian (1995) identify restrictions that are required to make the process non-Markovian.

One-factor models:

1. Absolute: $s(.) = s_0$,[13]

2. Linear Absolute: $s(.) = [s_0 + s_1(T\text{-}t)]$,

3. Square root: $s(.) = s_0 f(t,T)^{1/2}$,

4. Proportional: $s(.) = s_0 f(t,T)$,[14]

5. Linear proportional: $s(.) = [s_0 + s_1(T\text{-}t)]f(t,T)$.

Two-factor models:

1. Absolute: $s_1(.) = s_{1,}$

   $s_2(.) = s_2$.

2. Proportional: $s_1(.) = s_1 f(t,T)$,

   $s_2(.) = s_2 f(t,T)$.

2.3 Market Models

These models are consistent with market pricing practice for short term interest rates; hence, they are straightforward to calibrate using the Black (1976) formula for options on forward/futures contracts. For a particular tenor, t, market quoted forward rates are required to be log-normal. The tenor is fixed once and for all, since the requirement is that rates of only that tenor are log-normal. If $L(t,x)$ is the market quoted forward rate at time t for time $t+x$ of tenor t, then the process for the market quoted rate is required to be log-normal as follows:

$$dL(t,x) = m(t,x)dt + g(t,x)L(t,x)dz_t \tag{6}$$

where $g(t,x)$ is a d-dimensional vector. BGM show that for this restriction to hold, the drift $m(t,x)$ must have the form

$$\frac{\partial}{\partial x}L(t,x) + L(t,x)g(t,x)s(t,x) + \frac{t\ L^2(t,x)}{1+t\ L(t,x)}\left|g(t,x)\right|^2 \tag{7}$$

where $s(t,x)$ is related to $g(t,x)$ by

---

[13] This form of volatility specification leads to the continuous-time version of the Ho-Lee model, with Gaussian interest rates.

[14] The HJM framework requires that the volatility functions be bounded. Hence this volatility function is capped at a sufficiently high level of f^*, such that there is no effect on prices.

$$s(t,x) = \begin{cases} 0, & 0 \le x \le t \\ \displaystyle\sum_{k=1}^{\left\lfloor \frac{x}{t} \right\rfloor} \frac{t\, L(t, x-kt)}{1 + t\, L(t, x-kt)} g(t, x-kt), & t \le x. \end{cases} \tag{8}$$

The BGM functions *g(t,x)* are calibrated to the observed Black implied volatilities using the following relation

$$s_i^2 = \frac{1}{t_{i-1} - t} \int_t^{t_{i-1}} \left| g(s, t_{i-1} - s) \right|^2 ds \tag{9}$$

Since the BGM models focus on market quoted instruments, there is no need for instantaneous rates, which are required in the other models.

3. Model estimation and implementation

The spot rate models (HW and BK) are implemented by constructing a recombining trinomial lattice for the short-term interest rate.[15] The current term structure is estimated from spot LIBOR rates and Eurodollar futures prices, as explained in Appendix A. The volatility parameter s and the mean-reversion parameter a are chosen so as to provide a "best fit" to the market prices of caps and floors, by minimizing the sum of squared residuals. The delta hedge ratios are computed using the quadratic approximation to the first derivative of the option price with respect to the short rate.

Forward rate models are implemented under the HJM framework, with specific volatility functions, to ensure that the interest-rate process is Markovian, i.e., path independent. Path-dependence renders the implementation of a term structure model infeasible, in general, except for special cases. These special cases include models in which interest rates are assumed to be normally distributed, or where the volatility structures meet certain conditions to remove path dependence.[16] Further,

[15] Details of the trinomial lattice construction methodology can be obtained from Hull and White (1994).

[16] Ritchken and Sankarasubramanian (1995) have identified the necessary and sufficient conditions on volatility structures that capture the path dependence by a single sufficient statistic (which represents the

from a computational perspective, option prices cannot, in general, be represented as simple solutions to partial differential equations, because of the need to model multiple points on the term structure; this leads to complex boundary conditions with multiple state variables. Due to these reasons, the models in this paper are implemented using discrete-time, non-recombining binomial trees, which are computationally efficient.

The forward rate process described above is arbitrage-free only in continuous time and, therefore, cannot be directly used to construct a discrete-time tree for the evolution of the forward curve. Therefore, the drift term in the forward rate process needs to be reformulated in discrete time.[17] The derivation of the drift term for the discrete-time approximation of the forward rate process for the one- and two-factor models is presented in Appendix B. The delta hedge ratios are again computed as before, using the quadratic approximation to the first derivative.

The BGM model is implemented using Monte Carlo simulation, in the interests of computational efficiency. We simulate 5000 different paths, using the initial given term structure, and use antithetic variance reduction techniques, to price all our options. Extensive robustness checks were done to ensure that the results were not sensitive to the number of simulated paths. The discretization of the forward rate process and its drift are taken from Hull (2000). The delta hedge ratios are computed using a central difference approximation.

4. Experimental design

We apply two broad sets of tests to the interest rate option data, static and dynamic, that need some explanation. The fundamental motivation for testing the static accuracy of interest rate models is to examine whether they are capable of predicting future option prices conditional on term structure information. This capability is best evaluated by the ex-ante price predictive ability of the model. It is important for valuation models to capture information from current observable market data, and translate them into accurate option prices.[18] Towards this end, in this study, models are calibrated

---

accumulated variance of the forward rate upto the current date), thus making the evolution of the term structure Markovian with respect to two state variables.

[17] The discrete time no-arbitrage conditions for the drift term have been adapted from Jarrow (1996) and Radhakrishnan (1998).

[18] This is especially true for Value-at-Risk systems, where the objective is to be able to accurate estimate option prices in the future, conditional on term structure information.

based on the market data on term structure parameters as well as option prices at the current date. Then, at a future date, the same model is used along with current term structure information to estimate option prices. The accuracy of the predicted option prices is judged by comparing them with the actual observed option prices. This is a "static" test of the models, in the sense that current option prices are used to calibrate the model and price the same option one period later. This test does not examine whether the changes in option prices and the ability to hedge them are in line with the model's predictions.

The dynamic tests of these models examine the fundamental assumption underlying the construction of arbitrage-free option pricing models, which is the possibility of replication of the option by a portfolio of other securities that are sensitive to the same source(s) of uncertainty.[19] A test of the dynamic accuracy of these models can be constructed by examining the accuracy of local replication portfolios. This test is conducted by first constructing a hedge based on a given model, and then examining how the hedge performs over a small time interval subsequently. An accurate model to hedge interest rate exposures must produce price changes similar to those observed in the market, conditional on the changes of its state variables. Hence, the hedging tests are indicative of the extent to which the term structure models capture the future movements in the yield curve, i.e., the dynamics of the term structure. In principle, it is possible for a model to perform well in static tests and yet fail in dynamic tests, since the two types of tests are measuring different attributes of the model.

In arbitrage-free term structure models, the input parameters are allowed to change over time. The parameter vector is re-estimated each time the option prices are observed in order to fit a snapshot of market observables. This procedure is more permissive than the one dictated by the assumption that parameters are either constant or time-dependent in a deterministic way. It allows parameters to behave like pseudo-stochastic variables, despite not being assumed as such in the formulation of the model's stochastic structure. In this paper, we examine the "local" accuracy of term structure models; hence, it is not necessary to impose any restrictions on the model parameters.

In addition to the valuation performance measures, three other criteria are used to assess these models:

[19] With continuous trading and continuous state variable sample paths, the only sensitivities that matter for hedging are the deltas, since with continuous re-balancing, higher order sensitivities need not be

1) the stability of the parameters and the model performance over time,
2) the presence of systematic biases in the pricing and hedging errors, and
3) the relative complexity and difficulty in estimating the models, including numerical efficiency.

4.1 Hedging interest rate caps and floors

Since caplets and floorlets are essentially options on the forward interest rate, they can be hedged with appropriate positions in the LIBOR forward market. In practice, they are most commonly hedged using the short term interest rate futures contract, the Eurocurrency futures contract, e.g. Eurodollar futures, due to the liquidity of the futures market, as well as availability of contracts up to a maturity of 10 years, in increments of 3 months. Strictly speaking, interest rate forward contracts are similar to, but not exactly the same as interest rate futures contracts. The difference between the two is due to the negative convexity of the forward contract.[20] This convexity difference affects the computed hedge ratio. The price of an interest rate futures contract on the expiration date is defined as 100 minus the spot interest rate on that date. Hence, a short position in a caplet (floorlet) can be hedged by going short (long) an appropriate number of futures contracts. The hedge position of the cap (floor) is the sum of the hedge positions for the individual caplets (floorlets) in the cap (floor), i.e., a series of futures contracts of the appropriate maturities, known as the futures strip.

The hedge position is constructed by computing the change in the price of the caplets for a unit (say 1 basis point) change in the forward rate, relative to the number of futures contracts of *appropriate* maturity that give the same change in value for the same unit change in the forward rate. This is the delta hedge ratio for the caplet. In the context of a particular term structure model, the delta can sometimes be defined in closed form. In this paper, the hedge ratios are calculated numerically as explained in Section 3. Various robustness checks are done to ensure that the discretization of the continuous time process does not materially affect the accuracy of the computed delta.

---

explicitly considered.

[20] For details, see Gupta and Subrahmanyam (2000).

A portfolio of a short position in a cap and a short position in an appropriate number of futures contracts is *locally* insensitive to changes in the forward rate, thus making it "delta-neutral." In theory, this delta-neutral hedge requires continuous rebalancing to reflect the changing market conditions. In practice, however, only discrete rebalancing is possible. The accuracy of a delta hedge depends on how well the model's assumptions match the actual movements in interest rates.

A caplet/floorlet can also be gamma-hedged in addition to being delta-hedged, by taking positions in a variety of LIBOR options. Gamma is the second derivative of the price of the caplet/floorlet with respect to a change in the interest rate. Gamma hedging refers to hedging against changes in the hedge ratio. Setting up a gamma-neutral hedge results in a lower hedge slippage over time. However, in principle, the accuracy of the gamma hedge in the context of a particular model could be different from the accuracy of the delta hedge within the same model. Therefore, the hedging performance of the models could be different if they were evaluated using both delta and gamma hedging, instead of just delta hedging. In this paper, term structure models are tested based only on their delta hedging effectiveness.

There is a conceptual issue relating to hedging that needs to be defined explicitly. The hedging for any interest rate derivative contract can be done either "within the model" or "outside the model." The "within the model" hedge neutralizes the exposure only to the model driving factor(s), which, in the case of a one-factor model, is the spot or the forward rate. The "outside the model" hedge is determined by calculating price changes with respect to exogenous shocks, which, *per se*, would have a virtually zero probability of occurrence within the model itself.[21] This "outside the model" procedure is, hence, conceptually internally inconsistent and inappropriate when testing one model against another.[22] The "within the model" hedge tests give very useful indications about the realism of the model itself. The discussion about "delta-hedging" in the previous paragraphs of this section deals only with "within the model" hedging. This is the type of hedging that is empirically examined in this paper.

4.2 Empirical design for testing static accuracy

[21] Examples of such exogenous shocks include jumps in the yield curve or in individual forward rates, changes in the volatilities of interest rates, etc. These are ruled out within the structure of most of the models examined in this paper.

[22] From a practitioner's viewpoint, this inconsistency may be less important than the actual performance of the hedge.

In order to evaluate the static accuracy of interest rate models, we measure the comparative performance of the models for pricing caps / floors by analyzing the magnitude of the out-of-sample cross-sectional pricing errors. As explained earlier, the spot rate models are first estimated using constant parameters so that the models fit the current term structure exactly, but the volatility structure only approximately (in a least squares sense). In the second estimation, the parameters in the spot rate models are made time-varying so that the models fit the volatility term structure exactly as well, by calibration to the observed prices of caps / floors. To examine the out-of-sample pricing performance of each model, the prices of interest rate caps and floors at date t_i are used to calibrate the term structure model and back out the requisite implied parameters. Using these implied parameter values and the current term structure at date t_{i+1}, the prices of caps and floors are computed at date t_{i+1}. The observed market price is then subtracted from the model-based price, to compute both the absolute pricing error and the percentage pricing error. This procedure is repeated for each cap and floor in the sample, to compute the average absolute and the average percentage pricing errors as well as their standard deviations. These steps are followed separately for each of the models being evaluated. Then, the absolute as well as percentage pricing errors are segmented by type of option (cap or floor), “moneyness” (in-the-money, at-the-money, and out-of-the-money) and maturity to test for systematic biases and patterns in the pricing errors. The coefficients of correlation between the pricing errors across the various models are also computed to examine how the models perform with respect to each other.

The cross-sectional pricing performance of the models is further examined using two different calibration methods. The objective of estimating pricing errors using alternative calibration methods is to test the robustness of the pricing results to estimation methodology. In the first one, the prices of ATM caps (of all maturities) are used to calibrate the term structure model.[23] This model is then used to price the away-from-the-money caps of all maturities on the *same* day. The same procedure is repeated for the floors. The model prices are compared with market prices, and the errors are analyzed in a manner similar to the one before. In the second method, the cap prices (of all strike rates and maturities) are generated using the models calibrated to floor prices (of all strike rates and maturities), and floor prices generated by calibrating the models to cap prices. These two tests are strictly cross-sectional in nature, as the prices of options on one day are used to

[23] The ATM cap is taken to be the one with the strike that is closest to ATM, since, in general, no fixed strike cap (or floor) will be exactly ATM.

price other options on the *same* day, while in the earlier procedure the prices of options on the previous day were used to estimate current option prices.

To study the possible systematic biases in the pricing performance of the models in more detail, the pricing errors for these models are analyzed. The market price of the cap/floor is regressed on its model forecast price to analyze the mispricing and identify the model that is most consistent with data.

4.3 Empirical design for testing dynamic accuracy

Tests for dynamic accuracy evaluate the comparative performance of the models in hedging caps/floors. This is implemented by analyzing the magnitude of the out-of-sample cross-sectional hedging errors. To examine the hedging performance of the models, the term structure models are calibrated at date t_i using the current prices of interest rate caps and floors, and the requisite parameters are backed out. Using the current term structure of interest rates as well as spot cap/floor prices, the delta-hedge portfolio is constructed. The hedge portfolio is constructed separately for caps and floors. Each of these hedge portfolios consists of caps (or floors) of the 4 maturities (2-, 3-, 4- and 5-years), across the 4 strike prices, and the appropriate number of Eurodollar futures contracts.

In constructing the delta hedge for a caplet/floorlet with interest rate futures contracts, the hedge position must take into account an institutional factor. Caps/floors are negotiated each trading date for various maturities; hence, the expiration dates of caplets could be any date in the month. In contrast, exchange-traded futures contracts expire on a particular date. The expiration dates of the futures contracts generally do not coincide with the expiration dates of the individual caplets (floorlets) in the cap (floor). Therefore, it is necessary to create a "synthetic" futures contract whose expiration date coincides with that of a particular caplet/floorlet, by combining two futures contracts with maturity dates on either side of the expiration date of the caplet/floorlet being hedged. In other words, we form a synthetic position in a hypothetical futures contract expiring on the caplet/floorlet expiration date, by interpolating between the two adjacent contracts

Using this hedge portfolio, the hedging error is computed at date t_{i+k}, to reflect a k-day rebalancing interval. The hedging error corresponds to the change in the value of the hedge portfolio over these *k* days. In order to test for the effect of the rebalancing interval, the hedging errors are computed using a five-day and a twenty-day rebalancing interval.[24] In both cases, the procedure is repeated for each model, and the hedging errors are analyzed.[25]

5. Data

The data for this study consists of daily prices of U.S. dollar (USD) caps and floors, for a ten-month period (March 1 – December 31, 1998), i.e. 219 trading days, across four different strike rates (6.5%, 7%, 7.5%, 8% for caps, and 5%, 5.5%, 6%, 6.5% for floors) and four maturities (2-, 3-, 4-, and 5-year).[26] These data were obtained from Bloomberg Financial Markets.

Table 1 presents descriptive statistics of the data set. The prices of the contracts are expressed in basis points, i.e., a price of 1bp implies that the price of the contract for a notional principal of $10,000 is $1. The average, minimum and maximum price of the respective contracts over the sample period are reported in this table. The table indicates that the prices of both caps and floors increase, on average, with maturity. The prices of caps (floors) decrease (increase) with the strike rate.

It should be noted that our sample period witnessed considerable volatility in the global fixed income markets. Several major events triggered by the Russian default and the Long Term Capital Management (LTCM) crisis jolted the fixed income cash and derivatives markets. Hence, the dollar cap and floor markets experienced greater variation in prices than usual. This is fortuitous since it implies that the empirical tests of the various models are that much more stringent and, as a result, our conclusions are likely to be robust.

---

[24] A five-day rebalancing interval corresponds to weekly portfolio rebalancing, while a twenty-day rebalancing interval approximates monthly rebalancing. The results using daily rebalancing are not reported in the paper as there was very little hedge slippage over one trading day, thereby leading to almost perfect hedging using any model. Longer term rebalancing intervals provide a more stringent test of the extent to which the dynamics of the underlying interest rate are embedded in the model. The longer rebalancing intervals are in line with the spirit of capital adequacy regulations based on the guidelines of the Bank for International Settlements.

[25] The results reported in this paper are robust to the specific number of time steps in the discrete interest rate trees. Tests were done to study the differences in results by using larger number of time steps, and the differences were insignificant.

[26] Therefore, there are 218 days for which the model forecasts are compared with market prices.

Since interest rate caps and floors are contracts with specific maturity *periods* rather than specific maturity *dates*, a complication arises while doing the hedging tests. For these tests, we need the market prices of the original cap / floor contract that was hedged using futures. However, each day the reported prices of caps and floors refer to prices of *new* contracts of corresponding maturities, and not to the prices of the contracts quoted before. Hence, there is no *market price* series for any *individual* cap / floor contract. For example, consider a 5-year cap quoted at date t_i, which is also hedged at date t_i. To evaluate the performance of this hedge at date t_{i+1}, we need the price of the *same* cap at date t_{i+1}, i.e. at date t_{i+1}, we need the price of a cap expiring in 5 years *less one day*. However, the cap price that is observed at date t_{i+1} is the price of a *new* cap expiring in 5 years, not 5 years less one day. This data problem is not specific to just caps and floors – it is present for all OTC contracts that are fixed maturity rather than fixed maturity date contracts.

To overcome this problem, we construct a price series for each cap / floor contract, each day, until the expiration of the contract. The current term structure and the current term structure of volatilities (from the current prices of caps / floors) are used to price the original cap / floor contract each day. This price is used as a surrogate for the market price of the cap / floor contract on that particular day. This price is a model price, and not a real market price. However, the hedging performance tests are still useful in identifying models that can set up more accurate hedges for the cap / floor contracts. At the very least, the tests will evaluate models in terms of their internal consistency in terms of hedging performance.

6. Results

This section examines the results obtained for all models. The models are estimated each day using the current term structure of volatility from cap / floor prices.

6.1 Parameter stability

To examine the stability of the parameters of the estimated models, summary statistics for the estimated parameters are reported in table 2. The parameter estimates across models are not directly comparable for several reasons. First, the models use different factors (spot rates and forward rates), with some of them being two-factor models. Second, the drift and volatility functions differ in functional form. Third, the number of parameters estimated varies across models. However, the stability of these parameters can be inferred from the estimate of the coefficient of variation for each parameter.

Our results show that there is some variation in parameter estimates across time. By definition, the models posit that the drift and volatility parameters are constant. One explanation for this divergence from theory is that there is a second or third factor driving the evolution of rates, which is manifesting itself in the form of time-varying parameters. Possible candidates for the additional factor could be stochastic volatility, or a curvature factor. In our results, though the parameters vary over time, they are fairly stable. The coefficient of variation for most parameters is below 0.5, and for many parameters it is below 0.33.[27] The mean, standard deviation, coefficient of variation, minimum value and the maximum value of the parameters are reported in table 2. Comparable statistics are difficult to provide for the BGM model, since model estimation involves calibration of many volatility functions, not specific parameters, each day.

For the one-factor and the two-factor models, the parameter values are more stable for one-parameter models, while the coefficients of variation are significantly higher for the two-parameter models. In the case of spot rate models, the mean- reversion rate has a small absolute value and high standard error relative to the mean estimate, indicating that it is observed with significant error. In the forward rate models, the slope parameters for the linear absolute and linear proportional models have very high coefficients of variation and very small absolute values, making their estimates less reliable. These results indicate that adding more parameters to the model improves the ability of the model to fit prices, but significantly hampers the stability of the estimated model. This is also the reason why no model with more than two parameters was analyzed in this study. Therefore, from a practical perspective, the one-parameter one-factor models provide accurate, stable results as far as the model parameters are concerned.

6.2 Pricing performance

[27] The stability of the parameter estimates can be judged by looking at the coefficient of variation of the estimates over the sample period. A coefficient of variation below 0.5 indicates that the volatility of the estimate was less than half of the mean estimate; thus, the parameter was fairly stable over the sample period.

The tests for the comparative pricing performance of the models are implemented using the methodology described in section 4. The results for these tests are reported in tables 3, 4, 5, 6, and 7. These results are for out-of-sample fits of model-based prices to the observed market prices.[28]

The summary statistics of the forecast errors are presented in table 3. The table provides a first impression about the empirical quality of the models. The average absolute error is below 1 bp for caps, indicating a very small bias in the models. For floors, the error is close to 3 bp for the absolute and linear absolute forward rate models, while it is less than 1 bp for the other models. A similar pattern is observed in average percentage errors, which are less than 2% in most of the cases, indicating a very small bias. Since the bid-ask spread in these markets is of the order of 2 bp, the fit of the models is good.

The average absolute errors and the average absolute percentage errors display a clear pattern. The average absolute percentage errors are roughly similar for caps and floors. Within the class of one-factor models, the absolute errors are highest for the constant volatility forward rate model (3.5 bp for caps and 6.8 bp for floors) and lowest for the proportional (lognormal) forward rate model (1.2 bp for caps and 2.7 bp for floors). All the other models fall in between these models, in terms of prediction errors.[29] The two-factor models have marginally lower pricing errors as compared to the one-factor models that they nest. For example, the two-factor lognormal model has an average absolute error of 1.1 bp for caps and 2.4 bp for floors, as compared to 1.2 bp and 2.7 bp respectively for the one-factor lognormal model. Also, the spot rate models with time-varying parameters have considerably lower pricing errors for caps as well as floors, as compared to those for the models with constant parameters. Making the parameters time varying brings down the errors to almost the level of two-factor models. In this case, the time-varying parameters appear to be acting as "pseudo-factors." The one-factor BGM model works as well as the one-factor proportional volatility model. Perhaps, the one-factor lognormal structure that is common to both models is more important than other aspects of the two models.

[28] Note that these models use one or two parameters estimated out-of-sample to simultaneously generate 16 cap and 16 floor prices each day. In terms of the number of options, the models price 304 caplets (19 caplets for 4 maturities and 4 strikes each) and 304 floorlets (19 floorlets for 4 maturities and 4 strikes each) every day.

[29] The linear proportional model has a slightly lower average pricing error (2.5 bp) for floors. However, it is a two-parameter model, for which the parameter estimates are more volatile than those for the corresponding one-parameter model that it nests.

Table 4 presents the correlation between the pricing errors for the different models. The pricing errors for the models are computed by averaging the difference between the model prices and the observed market prices for all the caps/floors priced each day. The correlations are reported separately for caps and floors. There is a common component in the errors for all the models, which can be due to data noise, presence of other factors, etc. However, the correlations are higher within one-parameter and two-parameter models; this emphasizes the importance of the number of parameters in determining the behavior of the models. The correlations are also higher within the spot rate and the forward rate models, and within one-factor and two-factor models. Moreover, the correlations are slightly lower for floors as compared to caps. One possible reason for this result is the higher average price for floors, that results in larger absolute errors, and hence a lower correlation between them.

Tables 5 and 6 present the absolute and percentage errors for the caps/floors for all the models, for the cross-sectional tests using different calibration methods. For results in table 5, the models are first calibrated using ATM cap/floor prices, and then the ITM and OTM cap/floor prices are estimated. The absolute and percentage errors in this case are lower than those in table 3, where the models are calibrated using cap/floor prices from the previous day. For this calibration, the one-factor BGM model has the lowest pricing errors, while the constant volatility Gaussian model has the highest error. The proportional volatility models have low pricing error, but they are out-performed by the BGM model. Again, the spot rate models with time-varying parameters have much lower pricing errors. The two-factor models have marginally lower pricing errors than the one-factor models that they nest. The pricing errors are lowered further in table 6, where the models are calibrated using caps to estimate floor prices, and using floors to estimate cap prices. Across models, the pattern of errors is similar to the previous table. The one-factor BGM model outperforms all other models. The lognormal forward rate model provides fairly accurate pricing performance, but not the most accurate in these two calibrations. However, it should be noted that the BGM model is designed to fit contemporaneous cap prices exactly. Hence, in these two alternative calibrations, it performs better that the other models, since these tests, strictly speaking, are not out-of-sample - some of the option prices are being used to price the rest of the options the *same* day. The magnitudes of the pricing errors from the cross-sectional tests reinforce the conclusion that two-factor models are only marginally better than one-factor models for pricing these options. The results from the two alternative calibration methods for the models reaffirm that the pricing results reported in table 3 are robust to changes in model calibration methods. They also show that calibrating models to current option prices (as done in tables 5 and 6) and to a full range of strike rates (done only in table 6) results in more accurate pricing performance.

Figures 1 and 2 plot the percentage errors for the models, as a function of their strikes. All the models tend to overprice short-dated caps / floors and under-price long-dated ones. However, the over- and under-pricing patterns are different for one-parameter and two-parameter models. The one-parameter models tend to compensate the over-pricing of short-dated options by under-pricing long-dated options. The two-parameter models display a slight hump at the 3 yr maturity stage. They overprice medium-term caps / floors more than the short-term ones, and then compensate by under-pricing the long-dated caps / floors. In terms of fitting errors, the two-parameter models are a marginally better fit than the one-parameter models that they nest.

To study the systematic biases in more detail, the following cross-sectional regression model is estimated for caps and floors separately:

$$(\textit{Market Price})_t = b_0 + b_1\,(\textit{Model Forecast Price})_t + e_t \qquad (8)$$

The results of this estimation are presented in table 7. The objective of this estimation is to identify which model is most consistent with the data. The slope coefficients (b_1) are significantly different from zero and insignificantly different from one for all the models, with a very high R-square value, which shows that the average prediction error in the models is quite small. Also, the b_1 coefficient is slightly greater than one for floors, and slightly smaller than one for caps, for most of the models. Similarly, the b_0 coefficient is negative for floors and positive for caps, across all models. Thus, the models tend to overprice floors and underprice caps, which is consistent with the results reported earlier in this section. The spot rate models with time-varying parameters show slightly different results - they tend to overprice options. In the time-varying implementations, caps are being underpriced less, while floors are being overpriced more.

More significantly, the patterns of mispricing display a clear skew across strike rates, for all maturities. All the models tend to over-price in-the-money (low strike) caps and underprice out-of-the-money (high strike) caps. In the case of floors, the models underprice out-of-the-money (low strike) and overprice in-the-money (high strike). These patterns are consistent across all maturities. The skew is the greatest for the constant volatility (Ho-Lee Gaussian model) and the least for the proportional volatility models (one-factor and two-factor lognormal models). For the square root volatility model, in which the distribution of the underlying rate is non-central chi-square (which is less skewed than lognormal), the extent of skew in the pricing errors is also in between the Gaussian and the lognormal models. These patterns are similar for caps and floors, and are consistent across spot rate and forward rate models, as well as one-factor and two-factor models.

This negative skew in the pricing errors is consistent with the hypothesis that fatter right tails in the distribution of the underlying interest rate would lead to under-pricing in out-of-the-money caps and floors. The results indicate that the risk-neutral distribution of the underlying interest rate has a thinner left tail and a fatter right tail than the assumed distribution for any of these models. The partial correction of the skew by the lognormal model suggests that a skew greater than that in the lognormal distribution may help to predict away-from-the-money cap and floor prices better.

A comparison of the results for the one-factor models with those for the two-factor models shows that fitting the skew in the distribution of the underlying interest rate improves the static performance of the model more than by introducing another stochastic factor in the model. For example, the average pricing error for the one-factor lognormal model (1.2 bp for caps and 2.7 bp for floors) is much less than the average pricing errors for the two-factor Gaussian model (2.6 bp for caps and 5.0 bp for floors).

6.3 Hedging performance

The tests for the comparative dynamic accuracy of the models are conducted using the methodology described in section 4.3. The results for this analysis are presented in table 8. The accuracy of hedging, and hence the accuracy of replication of the interest rate options, differs significantly across term structure models. The average percentage hedging errors reported in table 8 show that 2-factor models perform significantly better than one-factor models in hedging interest rate risk in caps and floors. The difference is more significant for longer rebalancing intervals. With a 5-day rebalancing interval, most one-factor model hedges result in an average percentage error of about 0.5% of the hedge portfolio value in caps, and about 0.5%-0.8% in floors. In the case of two-factor models, the 5-day average percentage error is reduced to less than 0.2%. With a 20-day rebalancing interval, the average percentage hedging error reduces from 1.6%-3% for various one-factor models to 0.5%-0.7% for the two-factor models. Interestingly, the hedging results for the time-varying implementation of the spot rate models are very different from the pricing results - making the parameters time-varying actually leads to consistently *larger* hedging errors, indicating that the stability of model parameter estimation is important for accurate hedging performance. The hedging errors are evidence of the overall effectiveness of the interest rate hedges created by the models over time. Hence, the hedging performance reflects the dynamic accuracy of the various term structure models.

Within the class of one-factor and two-factor models, the hedging errors do depict the trend observed in the pricing errors, of a higher skew in the underlying distribution leading to smaller errors. For example, for the 5-day rebalancing interval, the average percentage error for caps goes down from 0.68% for the Gaussian one-factor forward rate model to 0.33% for the lognormal one-factor forward rate model. Similarly, for the 20-day rebalancing interval, the error goes down from 2.44% to 1.62%, respectively. However, adding a second stochastic factor leads to a much larger reduction in the hedging errors. This result is different from the pricing results where fitting the skew correctly dominated the introduction of a second stochastic factor. The Gaussian two-factor forward rate model has an average percentage error of 0.19% for 5-day rebalancing and 0.54% for 20-day rebalancing, which is significantly lower than those for the one-factor lognormal forward rate model.

In previous research, principal component analysis of interest rates changes reveals the various factors that drive the evolution of the term structure.[30] The first factor is interpreted as "level" factor capturing parallel shifts in the term structure, and has been shown to contribute about 92% of the overall explained variance of interest rate changes. The second factor, interpreted as a "twist" factor in the yield curve, incorporating changes in the slope of the term structure, has been shown to contribute another 7% of the overall explained variance of interest rate changes.[31] The results in this paper show that, for accurate hedging of interest rate caps and floors, it is not enough to correctly model just the first factor. Modeling the second factor allows the incorporation of expected twists in the yield curve while determining state variable sensitivities, thereby leading to more accurate hedging. This also constitutes evidence against claims in the literature, that correctly specified and calibrated one-factor models can replace multi-factor models for hedging purposes.[32]

---

[30] See, for example, Brown and Schaefer (1994) and Rebonato (1998).
[31] The third factor, interpreted as the "curvature" factor, incorporates changes in the curvature of the term structure, and explains most of the residual 1% variance of interest rate changes. This third factor may be important for pricing swaptions and bond options, but not for pricing interest rate caps and floors.
[32] See, for example, Hull and White (1990), and Buser, Hendershott and Sanders (1990).

7. Conclusions

A variety of models of interest rate dynamics have been proposed in the literature to value interest rate contingent claims. While there has been substantial theoretical research on models to value these claims, their empirical validity has not been tested with equal rigor. This paper presents extensive empirical tests of the static and dynamic accuracy of term structure models in the interest rate cap and floor markets. The paper also examines, probably for the first time in the literature, actual price data for caps and floors across strike rates, with maturities extending out to 5 years.

Alternative one-factor and two factor models are examined based on the accuracy of their out-of-sample price prediction, and their ability to hedge caps and floors. Within the class of one-factor models, two spot rate, five forward rate, and one market model specifications are analyzed. For two-factor models, two forward rate specifications are examined. Overall, in terms of the out-of-sample static tests, the one-factor lognormal (proportional volatility) forward rate model is found to outperform the other competing one-factor models in pricing accuracy. The estimated parameters of this model are more stable than those for corresponding two-parameter models, indicating that one-parameter models result in more robust estimation. In contrast, the pricing errors allowing for time-varying implementation of the one-factor models are at the level of those for the two-factor models: the time-varying parameters appear to be acting as "pseudo-factors." However, making the parameters time-varying actually leads to consistently *larger* hedging errors, indicating that the stability of model parameter estimation is important for accurate hedging performance. The one-factor BGM model also provides accurate pricing results, but outperforms the lognormal model only in tests which are not strictly out-of-sample.

More significantly, the lognormal assumption in the distribution of the underlying forward rate leads to a smaller "skew" in pricing errors across strike rates, as compared to the errors obtained by using a Gaussian interest rate process. The pricing accuracy of two-factor models is found to be only marginally better than the corresponding one-factor models that they nest. Therefore, the results show that a positive skew in the distribution of the underlying rate helps to explain away-from-the-money cap and floor prices more accurately, while the introduction of a second stochastic factor has only a marginal impact on pricing caps and floor.

On the other hand, the tests for dynamic accuracy of these models show that two-factor models are more effective in hedging the interest rate risk in caps and floors. While fitting the skew improves hedging performance marginally, introducing a second stochastic factor in the term structure model leads to significantly more accurate hedging. The one-factor BGM model provides hedging accuracy similar to the one-factor lognormal forward rate model, perhaps due to the common lognormal structure, but is outperformed by two-factor models. The two factor models allow a better representation of the dynamic evolution of the yield curve, by incorporating expected changes in the slope of the term structure. Since the interest rate dynamics embedded in two-factor models is closer to the one driving the actual economic environment, as compared to one-factor models, they are more accurate in hedging interest rate caps and floors. This result is also evidence against claims in the literature that correctly specified and calibrated one-factor models could replace multi-factor models for hedging.

So what are the implications of these results for the pricing and hedging of caps and floors in particular, and interest rate contingent claims in general? For interest rate caps and floors, one-factor lognormal and BGM models have been found to be sufficiently accurate in pricing performance. However, even for these plain-vanilla options, there is a need to use two-factor models for accurate hedging. Therefore, for consistent pricing and hedging within a book, even for plain-vanilla options like caps and floors, there is evidence that strongly suggests using two-factor models, over and above fitting the skew in the underlying interest rate distribution. Whether there is need for a third factor driving the term structure is still an open question for research.[33] Introducing more stochastic factors in the model makes computations more time consuming, so there is a trade-off between the cost of implementing a model and the stability of the model parameters, on the one hand, and its accuracy, on the other. However, for consistent pricing and hedging of the interest rate exposures of more complicated interest rate contingent claims like swaptions and yield spread options, there may be significant benefits to using term structure models with three or more factors. We defer these issues to be explored in future research.

[33] Litterman and Scheinkman (1991) report that the third factor, modeling changes in the curvature of the term structure, is important in explaining price changes.

Appendix A

Estimation of the current term structure

The current LIBOR term structure is estimated using spot LIBORs and Eurodollar futures prices. Theoretically, market swap rates can also be used along with spot LIBORs to estimate the LIBOR term structure. However, swap rates are available only for maturities of 2, 3, 4, 5, 7, and 10 years, while Eurodollar futures prices are available for maturities upto 10 years in increments of three months, which allows the computation of LIBOR zero rates with much higher accuracy. Moreover, Eurodollar futures contracts are extremely liquid with very high trading volumes and open interest. Hence, they are likely to reflect the best available information about the term structure of interest rates.

The spot market data are used to accurately define the curvature of the LIBOR yield curve, going out to the first futures expiration date (0-3 months, depending on the date). Beyond that date, Eurodollar futures prices are used to estimate the yield curve going out to 10 years. The yield curve thus obtained is then corrected for convexity. It is well known that, in the presence of stochastic interest rates, the implied forward rates are lower than futures rates, due to convexity in the payoffs of forward contracts.[34] Hence, the convexity adjustments are estimated for each futures contract maturity, and then subtracted from the futures yield curve to obtain the convexity-corrected LIBOR zero curve. The convexity adjustments are computed using the Hull-White one-factor model, for which the parameters are estimated using current option prices. The same convexity adjustments (and hence, the same yield curve) are used for testing all the models.[35] The cubic spline interpolation method is used to define the complete shape of the yield curve as a smooth function of maturity.

[34] See Gupta and Subrahmanyam (2000) for a detailed description of convexity adjustments, and the methods that can be used to estimate them.

[35] The convexity adjustments are fairly invariant to the model used to estimate them, as shown in Gupta and Subrahmanyam (2000). Therefore, the use of the Hull-White model to estimate the convexity adjustment for constructing the yield curve, across models, is unlikely to make a difference.

Appendix B

Derivation of the drift term for the discrete-time approximation of the forward rate processes

The forward rate process in the HJM framework is arbitrage free only in continuous time. Hence, for discrete time implementations of the model, the drift term for the process needs to be reformulated, for the one- and two-factor models.

Discretization of a one-factor process leads to two branches at each node of the tree. In the discrete economy, let h_n be the time step from time t_n to t_{n+1}. Given that the forward rate process is in state s_t at time t_n, it can be in one of the following two states (up or down) at time t_{n+1}:

$$f(t_{n+1},T;s_{t_{n+1}}) = \begin{cases} f(t_n,T:s_{t_n}) + m(t_n,T:s_{t_n})h_n + s(t_n,T:s_{t_n})\sqrt{h_n} & \text{if } s_{t_{n+1}} = s_{t_n}u \text{ w.p. } 1/2 \\ f(t_n,T:s_{t_n}) + m(t_n,T:s_{t_n})h_n - s(t_n,T:s_{t_n})\sqrt{h_n} & \text{if } s_{t_{n+1}} = s_{t_n}d \text{ w.p. } 1/2 \end{cases}$$

…….(C.1)

where *m(.)* is the drift and *s(.)* is the volatility function of the forward rate process *f*. The maturity of the forward rate, *T*, can take on any value between *t* and the maximum maturity assumed in the term structure to generate as many forward rates as desired, within the constraints of computational limitations.

In this framework, discount bond prices are given by

$$P(t_n,T) = \exp\left(\sum_{j=n}^{T-1} - f(t_n,t_j)h_j\right) \tag{C.2}$$

These discount bond prices evolve in the following manner:

$$P(t_{n+1},T;s_{t_{n+1}}) = \begin{cases} P(t_n,T;s_{t_n})\,pu(t_n,T;s_{t_n}) & \text{if } s_{t_{n+1}} = s_{t_n}u \text{ w.p. } 1/2 \\ P(t_n,T;s_{t_n})\,pd(t_n,T;s_{t_n}) & \text{if } s_{t_{n+1}} = s_{t_n}d \text{ w.p. } 1/2 \end{cases} \tag{C.3}$$

Given the above processes for the forward rates and the discount bond prices, *pu(.)* and *pd(.)* can be represented in terms of the *m(.)* and *s(.)* functions.

Using the money market account as the numeraire implies that all bond prices grow at the riskless rate, $f(t_n,t_n)$. Therefore, the martingale condition applied to the discrete framework requires that

$$P(t_n,T) = E_t\left[P(t_n,t_{n+1}).P(t_{n+1},T)\right] \tag{C.4}$$

i.e.,

$$P(t_n,T;s_{t_n})\, f(t_n,t_n;s_{t_n}) = P(t_n,T;s_{t_n})\left[\tfrac{1}{2}\, pu(t_n,T;s_{t_n}) + \tfrac{1}{2}\, pd(t_n,T;s_{t_n})\right] \tag{C.5}$$

This is a system of N equations, where N is the number of forward rates at time t_n, that can be solved recursively to get the expression of the drift term, $m(.)$, in each of the N forward rate processes:

$$m(t_n,T;s_{t_n}) = \frac{1}{h_n h_{T-1}} \ln\left[\frac{\cosh\left(\sum_{j=n+1}^{T} s\,(t_n,t_j;s_{t_n})h_j\sqrt{h_n}\right)}{\cosh\left(\sum_{j=n+1}^{T-1} s\,(t_n,t_j;s_{t_n})h_j\sqrt{h_n}\right)}\right] \tag{C.6}$$

Using a specific functional form for the volatility function $s(.)$, the drift from the equation above and the forward rate process evolution, the HJM interest rate tree can now be constructed.

For the two-factor process, discretization requires three branches at each node of the tree. The forward rate process is represented in a manner similar to the one-factor case, as follows:

$$f(t_{n+1},T;s_{t_{n+1}}) = \begin{cases} f(t_n,T:s_{t_n}) + m(t_n,T:s_{t_n})h_n - s_1(t_n,T:s_{t_n})\sqrt{h_n} + \sqrt{2}\, s_2(t_n,T:s_{t_n})\sqrt{h_n} \\ \qquad \text{if } s_{t_{n+1}} = s_{t_n}u \text{ w.p. } 1/4 \\ f(t_n,T:s_{t_n}) + m(t_n,T:s_{t_n})h_n - s_1(t_n,T:s_{t_n})\sqrt{h_n} - \sqrt{2}\, s_2(t_n,T:s_{t_n})\sqrt{h_n} \\ \qquad \text{if } s_{t_{n+1}} = s_{t_n}m \text{ w.p. } 1/4 \\ f(t_n,T:s_{t_n}) + m(t_n,T:s_{t_n})h_n + s_1(t_n,T:s_{t_n})\sqrt{h_n} \\ \qquad \text{if } s_{t_{n+1}} = s_{t_n}d \text{ w.p. } 1/2 \end{cases}$$

……..(C.7)

Using the money market account as the numeraire and applying the martingale condition, the drift function in discrete time is given by

$$
m(t_n,T;s_{t_n}) = \frac{1}{h_n h_{T-1}} \ln \left[\begin{array}{l} \frac{1}{2}\exp\left(-\sum_{j=n+1}^{T} s_1(t_n,t_j;s_{t_n}) h_j \sqrt{h_n} \right) \\ + \frac{1}{2}\exp\left(-\sum_{j=n+1}^{T} s_1(t_n,t_j;s_{t_n}) h_j \sqrt{h_n} \right) \left\{ \begin{array}{l} \frac{1}{2}\exp\left(\sqrt{2} \sum_{j=n+1}^{T} s_2(t_n,t_j;s_{t_n}) h_j \sqrt{h_n} \right) \\ + \frac{1}{2}\exp\left(-\sqrt{2} \sum_{j=n+1}^{T} s_2(t_n,t_j;s_{t_n}) h_j \sqrt{h_n} \right) \end{array} \right\} \end{array} \right]
$$

$$
- \frac{1}{h_n h_{T-1}} \sum_{j=n+1}^{T-1} m(t_n,t_j;s_{t_n}) h_{j-1} h_n \qquad \text{……..(C.8)}
$$

As in the one-factor case, the HJM tree can now be constructed using any specific form for the volatility function.[36]

[36] The HJM tree is non-recombining, due to the non-Markovian nature of the forward rate process for most volatility structures. Hence, from a numerical implementation perspective, the exploding number of terminal nodes in the tree imposes a limit on the number of time steps that can be used for a general volatility structure. In the usual binomial tree, the burden on computer memory and computing power is enormous since each node has to carry the values of the entire forward rate vector. Therefore, in this paper, a recursive algorithm proposed by Das (1998) is used. This algorithm eliminates the need for storing the entire forward rate tree in the memory, by following each sample path to its conclusion in a recursive manner. This frees up memory space, potentially allowing a relatively large number of time steps to be used, within the constraints of computing time, and also speeds up computation. See Das (1998) for details of the recursive algorithm.

References

Amin, Kaushik I., and Andrew Morton, 1994, Implied volatility functions in arbitrage-free term structure models, *Journal of Financial Economics* 35, 141-180.

Amin, Kaushik I., and James N. Bodurtha, Jr., 1995, Discrete-time valuation of American options with stochastic interest rates, *Review of Financial Studies* 8(1), 193-234.

Andersen, Lief, 1999, A simple approach to the pricing of bermudan swaptions in the multi-factor Libor market model, *Working Paper, General Re Financial Products.*

Black, Fisher, 1976, The pricing of commodity contracts, *Journal of Financial Economics* 3, 167-179.

Black, Fisher, Emanuel Derman, and William Toy, 1990, A one-factor model of interest rates and its application to Treasury bond options, *Financial Analysts Journal* 46, 33-39.

Black, Fisher, and Piotr Karasinski, 1991, Bond and option pricing when short rates are lognormal, *Financial Analysts Journal* 47, 52-29.

Brennan, Michael, and Eduardo Schwartz, 1979, A continuous time approach to the pricing of bonds, *Journal of Banking and Finance* 3, 133-155.

Brennan, Michael, and Eduardo Schwartz, 1982, An equilibrium model of bond pricing and a test of market efficiency, *Journal of Financial and Quantitative Analysis* 17, 301-329.

Brown, Roger H., and Stephen M. Schaefer, 1994, Why do long-term forward rates (almost always) slope downwards?, *Working Paper, London Business School.*

Bühler, Wolfgang, Marliese Uhrig, Ulrich Walter, and Thomas Weber, 1999, An empirical comparison of forward- and spot-rate models for valuing interest-rate options, *Journal of Finance* 54, 269-305.

Buser, Stephen A., Patrick Hendershott, and Anthony Sanders, 1990, Determinants of the value of the call options on default-free bonds, *Journal of Business* 63:1, 33-50.

Canabarro, Eduardo, 1995, Comparing the dynamic accuracy of yield-curve-based interest rate contingent claim pricing models, *Journal of Financial Engineering* 2, 365-401.

Chan, K.C., G. Andrew Karolyi, Francis A. Longstaff, and Anthony B. Saunders, 1992, An empirical comparison of alternative models of the short-term interest rate, *Journal of Finance* 47, 1209-1227.

Cox, John C., Jonathan E. Ingersoll, and Stephen A. Ross, 1985, A theory of the term structure of interest rates, *Econometrica* 53, 385-408.

Das, Sanjiv R., 1998, On the recursive implementation of term-structure models, *Working Paper, Harvard Business School.*

De Jong, Frank, Joost Driessen, and Antoon Pelsser, 1999, Libor and swap market models for the pricing of interest rate derivatives: An empirical analysis, *Working Paper.*

Driessen, Joost, Pieter Klaasen, and Bertrand Melenberg, 2000, The performance of multi-factor term structure models for pricing and hedging caps and swaptions, *Working Paper.*

Duffie, Darrell, and Rui Kan, 1996, A yield-factor model of interest rates, *Mathematical Finance* 64, 379-406.

Flesaker, Bjorn, 1993, Testing the Heath-Jarrow-Morton / Ho-Lee model of interest-rate contingent claims pricing, *Journal of Financial and Quantitative Analysis* 28, 483-495.

Gupta, Anurag, and Marti G. Subrahmanyam, 2000, An empirical examination of the convexity bias in the pricing of interest rate swaps, *Journal of Financial Economics,* 55, 239-279.

Heath, David, Robert A. Jarrow, and Andrew Morton, 1990a, Bond pricing and the term structure of interest rates: A discrete time approximation, *Journal of Financial and Quantitative Analysis* 25, 419-440.

Heath, David, Robert A. Jarrow, and Andrew Morton, 1990b, Bond pricing and the term structure of interest rates: A new methodology for contingent claims valuation, *Econometrica* 60, 77-105.

Ho, Thomas S. Y., and Sang-Bin Lee, 1986, Term structure movements and pricing interest rate contingent claims, *Journal of Finance* 41, 1011-1029.

Hull, John, and Alan White, 1990, Pricing interest rate derivative securities, *Review of Financial Studies* 3, 573-592.

Hull, John, and Alan White, 1993, One-factor interest-rate models and the valuation of interest-rate derivative securities, *Journal of Financial and Quantitative Analysis* 28, 235-254.

Hull, John, and Alan White, 1994, Numerical procedures for implementing term structure models: single-factor models, *Journal of Derivatives* 2, 7-16.

Hull, John, and Alan White, 1995, "A note on the models of Hull and White for pricing options on the term structure": Response, *Journal of Fixed Income* 5(2), 97-102.

Hull, John, and Alan White, 1996, Using Hull-White interest rate trees, *Journal of Derivatives* Spring, 26-36.

Hull, John, and Alan White, 1999, Forward rate volatilities, swap rate volatilities, and the implementation of the LIBOR market model, *Working Paper, University of Toronto.*

Jarrow, Robert, 1996, *Modeling Fixed Income Securities and Interest Rate Options* (McGraw-Hill).

Litterman, Robert and José Scheinkman, 1991, Common factors affecting bond returns, *Journal of Fixed Income* June, 54-61

Longstaff, Francis A., and Eduardo Schwartz, 1992, Interest-rate volatility and the term structure: A two-factor general equilibrium model, *Journal of Finance* 47, 1259-1282.

Longstaff, Francis A., and Eduardo Schwartz, 1993, Implementation of the Longstaff-Schwartz interest-rate model, *Journal of Fixed Income* 3, 7-14.

Longstaff, Francis A., and Eduardo Schwartz, 1995, A simple approach to valuing risky fixed and floating rate debt, *Journal of Finance* 50, 789-819.

Longstaff, Francis A., Pedro Santa-Clara, and Eduardo Schwartz, 2000, The relative valuation of caps and swaptions: Theory and empirical evidence, *Working Paper, UCLA*.

Miltersen, Kristian R., Klaus Sandmann and Dieter Sondermann, 1997, Closed Form Solutions for Term Structure Derivatives with Log-Normal Interest Rates, *Journal of Finance, 409-430*.

Moraleda, Juan M. and Antoon Pelsser, 2000, Forward versus spot interest rate models of the term structure: An empirical comparison, *The Journal of Derivatives* Spring, 9-21.

Pearson, Neil, and Tong-Sheng Sun, 1994, Exploiting the conditional density in estimating the term structure: An application to the Cox, Ingersoll, and Ross model, *Journal of Finance* 49, 1279-1304.

Peterson, Sandra, Richard C. Stapleton, and Marti Subrahmanyam, 1999, A two-factor lognormal model of the term structure and the valuation of American-style options on bonds, *Working Paper, New York University*.

Radhakrishnan, A. R., 1998, An empirical study of the convergence properties of non-recombining HJM forward rate tree in pricing interest rate derivatives, *Working Paper, New York University*.

Rebonato, Riccardo, 1998, *Interest-Rate Option Models* 2nd Ed. (John Wiley & Sons, Chichester, England).

Ritchken, Peter and Iyuan Chuang, 1999, Interest rate option pricing with volatility humps, *Review of Derivatives Research* 3, 237-262.

Ritchken, Peter, and L. Sankarasubramanian, 1995, Volatility structures of forward rates and the dynamics of the term structure, *Mathematical Finance* 5, 55-72.

Schaefer, M. Stephen, and Eduardo S. Schwartz, 1984, A two-factor model of the term structure – An approximate analytical solution, *Journal of Financial and Quantitative Analysis* 19, 413-424.

Stapleton, Richard C., and Marti Subrahmanyam, 1999, A two-factor stochastic model of interest rate futures prices, *Working Paper, New York University*.

Vasicek, Oldrich, 1977, An equilibrium characterization of the term structure, *Journal of Financial Economics* 5, 177-188.

Table 1

This table presents descriptive statistics of the data set used in this paper. The data consists of cap and floor prices across 4 different maturities (2-, 3-, 4-, and 5-year) and across 4 different strike rates, for each maturity (6.5%, 7%, 7.5%, and 8% for caps and 5%, 5.5%, 6%, and 6.5% for floors). The sample period consists of 219 trading days of daily data, from March 1 to December 31, 1998. The prices of the contracts are expressed in basis points, i.e., a price of 1bp implies that the price of the contract for a notional principal of $10,000 is $1. The average, minimum and maximum price of the respective contracts over the sample period are reported in this table.

	6.5% Caps				7% Caps			
	2 yr	3 yr	4 yr	5 yr	2 yr	3 yr	4 yr	5 yr
Mean	16	37	72	117	8	22	47	82
Min	4	13	32	57	2	8	21	42
Max	33	64	109	164	18	38	74	120
	7.5% Caps				8% Caps			
	2 yr	3 yr	4 yr	5 yr	2 yr	3 yr	4 yr	5 yr
Mean	4	13	31	57	3	8	20	40
Min	2	3	12	29	1	2	8	21
Max	10	24	55	94	5	17	41	75
	5% Floors				5.5% Floors			
	2 yr	3 yr	4 yr	5 yr	2 yr	3 yr	4 yr	5 yr
Mean	37	132	163	197	67	186	234	284
Min	7	80	98	115	20	112	143	169
Max	129	267	328	385	190	359	445	523
	6% Floors				6.5% Floors			
	2 yr	3 yr	4 yr	5 yr	2 yr	3 yr	4 yr	5 yr
Mean	116	262	332	401	182	363	461	557
Min	51	166	213	254	106	251	322	385
Max	262	465	580	682	341	583	731	864

Table 2

This table presents summary statistics for the parameter estimates for the one-factor and two-factor spot rate and forward rate models tested in this paper. The summary statistics for each parameter are computed using daily parameter estimates over the sample period, March 1 - December 31, 1998. The models are estimated each day over the 219 day sample period, by calibrating them to the market prices of caps and floors across four different maturities (2-, 3-, 4-, and 5-year) and across four different strike rates for each maturity (6.5%, 7%, 7.5%, 8% for caps, and 5%, 5.5%, 6%, 6.5% for floors).

Model	Parameter	Mean	Min	Max	s.d.	c.v.
Spot Rate Models						
Hull and White	a	0.045	0	0.088	0.027	0.61
	σ	0.0109	0.0051	0.0172	0.0035	0.32
Black and Karasinski	a	0.055	0	0.097	0.025	0.45
	σ	0.194	0.131	0.284	0.056	0.29
Forward Rate Models – One Factor						
Absolute	σ_0	0.0113	0.0075	0.0214	0.0035	0.31
Linear Absolute	σ_0	0.0098	0.0031	0.018	0.0043	0.44
	σ_1	0.0007	-0.0029	0.053	0.0018	2.6
Square Root	σ_0	0.0456	0.0273	0.0874	0.0105	0.23
Proportional	σ_0	0.1851	0.1169	0.2741	0.0407	0.22
Linear Proportional	σ_0	0.1759	0.0799	0.2632	0.0721	0.41
	σ_1	0.0053	-0.0005	0.0138	0.0037	0.70
Forward Rate Models –Two Factor						
Absolute	σ_1	0.0051	0.0021	0.0107	0.0023	0.45
	σ_2	0.0101	0.0039	0.0206	0.0047	0.47
Proportional	σ_1	0.0894	0.0513	0.1374	0.0277	0.31
	σ_2	0.1621	0.0906	0.2591	0.0438	0.27

Table 3

This table presents summary statistics for the forecast errors (in basis points and percentage terms) for the one-factor and two-factor spot rate and forward rate models analyzed in the paper. The average error is defined as the predicted model price minus the observed market price, averaged for the 32 caps and floors (4 strike rates each for caps and floors, for each of the 4 maturities) over the 219 days (March-December, 1998) for which the study was done. The average percentage error is defined as the (model price – market price) / market price, averaged in a similar way.

Model	Caps				Floors			
	Avg Error (bp)	Avg Abs Error (bp)	Avg % Error	Avg % Abs Error	Avg Error (bp)	Avg Abs Error (bp)	Avg % Error	Avg % Abs Error
Spot Rate Models								
Hull and White	-1.0	2.3	-1.8%	6.9%	1.0	4.6	-1.3%	5.3%
HW - time varying	-0.2	1.3	-0.9%	4.5%	2.5	3.9	0.5%	3.8%
Black & Karasinski	0.1	1.4	0%	4.3%	-0.1	3.0	-1.3%	3.1%
BK - time varying	0.4	1.1	0.7%	3.3%	0.2	2.5	-0.7%	2.4%
Forward Rate Models – One Factor								
Absolute	0.8	3.5	1.4%	10.1%	2.9	6.8	-0.3%	6.0%
Linear Absolute	0.1	2.3	-0.2%	6.9%	2.6	6.2	-0.7%	6.4%
Square Root	-1.2	1.7	-2.3%	4.9%	0.5	3.8	-1.0%	3.9%
Proportional	0.1	1.2	0%	4.0%	-0.1	2.7	-1.3%	2.9%
Linear Proportional	0.2	1.2	0.6%	3.9%	-0.1	2.5	-1.1%	2.7%
Forward Rate Models – Two Factor								
Absolute	0.8	2.6	1.4%	7.8%	2.3	5.0	-0.1%	4.6%
Proportional	0.05	1.1	0%	3.7%	-0.1	2.4	-1.1%	2.6%
Market Model - One Factor								
BGM	0.5	1.2	0.7%	3.9%	0.1	2.6	-1.2%	2.8%

Table 4

This table presents the correlation coefficients between the out-of-sample pricing errors for caps and floors for the models tested in this paper. There are a total of 219 observations for each model, corresponding to each day over the sample period, March 1 - December 31, 1998. The pricing error is defined as the model price minus the observed market price, averaged across the 32 caps and floors priced on that day (4 strike rates each for caps and floors, for each of the 4 maturities).

Model	HW	HW (II)	BK	BK (II)	Abs. (1-fac)	Linear Abs.	Square Root	Prop. (1-fac)	Linear Prop.	Abs. (2-fac)	Prop. (2-fac)	BGM (1-fac)
Caps												
HW	1											
HW (II)	0.94	1										
BK	0.87	0.83	1									
BK (II)	0.79	0.81	0.92	1								
Abs. (1-fac)	0.81	0.77	0.54	0.59	1							
Linear Abs.	0.69	0.72	0.49	0.61	0.71	1						
Square root	0.53	0.63	0.76	0.75	0.89	0.72	1					
Prop. (1-fac)	0.57	0.61	0.96	0.98	0.86	0.75	0.97	1				
Linear Prop.	0.42	0.50	0.92	0.94	0.63	0.96	0.75	0.69	1			
Abs. (2-fac)	0.49	0.48	0.43	0.56	0.82	0.71	0.54	0.55	0.48	1		
Prop. (2-fac)	0.41	0.47	0.65	0.69	0.63	0.58	0.46	0.92	0.81	0.77	1	
BGM (1-fac)	0.55	0.49	0.54	0.44	0.61	0.58	0.60	0.72	0.68	0.53	0.55	1
Floors												
HW	1											
HW (II)	0.95	1										
BK	0.84	0.82	1									
BK (II)	0.79	0.75	0.94	1								
Abs. (1-fac)	0.78	0.71	0.59	0.64	1							
Linear Abs.	0.65	0.66	0.45	0.51	0.69	1						
Square root	0.61	0.72	0.71	0.69	0.85	0.68	1					
Prop. (1-fac)	0.56	0.63	0.98	0.93	0.84	0.74	0.96	1				
Linear Prop.	0.49	0.44	0.91	0.90	0.59	0.94	0.69	0.62	1			
Abs. (2-fac)	0.47	0.46	0.49	0.48	0.79	0.74	0.51	0.57	0.44	1		
Prop. (2-fac)	0.44	0.51	0.59	0.55	0.63	0.61	0.41	0.89	0.75	0.72	1	
BGM (1-fac)	0.39	0.47	0.40	0.49	0.58	0.52	0.61	0.73	0.66	0.48	0.53	1

Table 5

This table presents summary statistics for the cross-sectional out-of-sample forecast errors (in basis points and percentage terms) for the one-factor and two-factor spot rate and forward rate models. The models are calibrated using the prices of ATM options (out of the 4 strike rates, the one that is closest to ATM). Then, the prices of the away-from-the-money (ITM and OTM) caps and floors are estimated using the models (for the 3 remaining strike rates). This is done for all maturities, and for caps and floors separately. The average error is defined as the predicted model price minus the observed market price, averaged for the 12 caps/floors (the 3 remaining strike rates for each of the 4 maturities) over the 219 days (March-December, 1998) for which the study was done. The average percentage error is defined as the (model price – market price)/market price, averaged in a similar way.

Model	Caps				Floors			
	Avg Error (bp)	Avg Abs Error (bp)	Avg % Error	Avg % Abs Error	Avg Error (bp)	Avg Abs Error (bp)	Avg % Error	Avg % Abs Error
Spot Rate Models								
Hull and White	-0.7	1.8	-1.3%	4.9%	0.8	3.3	-0.8%	3.7%
HW - time varying	0.1	1.1	0%	3.3%	1.9	3.0	0.1%	3.0%
Black & Karasinski	0.1	0.9	0.1%	2.6%	0	1.9	-0.8%	1.9%
BK - time varying	0.4	0.7	0.5%	1.9%	1.2	1.7	0.1%	1.5%
Forward Rate Models – One Factor								
Absolute	0.6	2.7	1.2%	7.8%	2.2	5.2	-0.2%	4.5%
Linear Absolute	0.1	1.8	0.1%	5.1%	2.0	4.5	-0.5%	4.7%
Square Root	-0.9	1.2	-1.7%	3.4%	0.4	3.0	-0.7%	2.9%
Proportional	0.1	0.9	0%	2.4%	-0.04	1.9	-0.8%	1.9%
Linear Proportional	-0.1	0.8	0.1%	2.4%	0.05	1.8	-0.6%	1.8%
Forward Rate Models – Two Factor								
Absolute	0.5	1.7	0.8%	5.0%	1.5	3.4	-0.2%	3.1%
Proportional	0.02	0.6	0%	1.9%	-0.06	1.5	-0.7%	1.6%
Market Model - One Factor								
BGM	0	0.5	0%	1.6%	0	0.9	0%	1.2%

Table 6

This table presents summary statistics for the cross-sectional out-of-sample forecast errors (in basis points and percentage terms) for the one-factor and two-factor spot rate and forward rate models. For pricing caps, the models are calibrated using the current prices of floors, and vice-versa. The average error is defined as the predicted model price minus the observed market price, averaged for the 16 caps or floors (4 strike rates for each of the 4 maturities) over the 219 days (March-December, 1998) for which the study was done. The average percentage error is defined as the (model price – market price)/market price, averaged in a similar way.

Model	Caps				Floors			
	Avg Error (bp)	Avg Abs Error (bp)	Avg % Error	Avg % Abs Error	Avg Error (bp)	Avg Abs Error (bp)	Avg % Error	Avg % Abs Error
Spot Rate Models								
Hull and White	-0.5	1.2	-0.9%	3.5%	0.6	2.3	-0.6%	2.5%
HW - time varying	0.2	0.8	0.1%	2.5%	1.3	2.0	0.1%	2.0%
Black & Karasinski	0.1	0.6	0.1%	1.8%	0	1.2	-0.5%	1.2%
BK - time varying	0.3	0.5	0.5%	1.3%	0.3	0.9	-0.2%	0.9%
Forward Rate Models – One Factor								
Absolute	0.5	2.0	0.9%	5.7%	1.6	3.7	-0.2%	3.2%
Linear Absolute	0.1	1.3	0.1%	3.8%	1.5	3.3	-0.4%	3.3%
Square Root	-0.6	0.8	-1.1%	2.3%	0.3	2.0	-0.5%	1.9%
Proportional	0.04	0.5	0%	1.5%	-0.02	1.2	-0.5%	1.2%
Linear Proportional	-0.1	0.5	0%	1.4%	0.03	1.1	-0.4%	1.1%
Forward Rate Models – Two Factor								
Absolute	0.4	1.2	0.6%	3.5%	0.9	2.0	-0.1%	1.9%
Proportional	0.02	0.4	0%	1.1%	-0.04	0.9	-0.4%	0.9%
Market Model - One Factor								
BGM	0	0.4	0%	0.9%	0	1.0	0%	0.8%

Table 7

This table presents results for model performance by estimating the following regression model for each of the one-factor and two-factor models examined in the paper:

$$(Market\ Price)_t = b_0 + b_1\ (Model\ Forecast\ Price)_t + e_t$$

The coefficients are presented with the standard errors in parenthesis for the slope coefficient. The model and market prices of the caps and floors are expressed in basis points, for the 219 daily observations during the sample period March-December, 1998. All the caps (6.5%, 7%, 7.5%, and 8% strike) and floors (5%, 5.5%, 6%, 6.5%) for each of the four maturities (2-, 3-, 4-, and 5-year) are used in the regression model to test for biases in model performance.

Model	Caps			Floors		
	b_0	b_1	R^2	b_0	b_1	R^2
Spot Rate Models						
Hull and White	2.538	0.971 (0.027)	0.978	-1.153	1.012 (0.019)	0.983
HW - time varying	1.107	0.988 (0.021)	0.991	-2.213	1.029 (0.017)	0.994
Black & Karasinski	0.083	1.002 (0.013)	0.994	-0.094	0.997 (0.008)	0.991
BK - time varying	-0.671	1.015 (0.012)	0.996	-1.379	1.021 (0.011)	0.997
Forward Rate Models: One-Factor						
Absolute	-0.094	1.007 (0.020)	0.977	-3.217	1.025 (0.022)	0.972
Linear Absolute	0.065	1.002 (0.018)	0.989	-2.439	1.019 (0.016)	0.979
Square Root	2.972	0.963 (0.029)	0.982	-0.328	1.002 (0.019)	0.984
Proportional	0.039	1.001 (0.014)	0.993	0.049	0.998 (0.009)	0.995
Linear Proportional	0.070	1.003 (0.011)	0.994	-0.055	0.997 (0.012)	0.995
Forward Rate Models: Two-Factor						
Absolute	-0.046	1.008 (0.013)	0.988	-2.057	1.017 (0.015)	0.980
Proportional	0.015	1.000 (0.006)	0.997	-0.028	0.999 (0.007)	0.998
Market Model - One Factor						
BGM	0.043	1.001 (0.007)	0.991	0.022	0.999 (0.005)	0.996

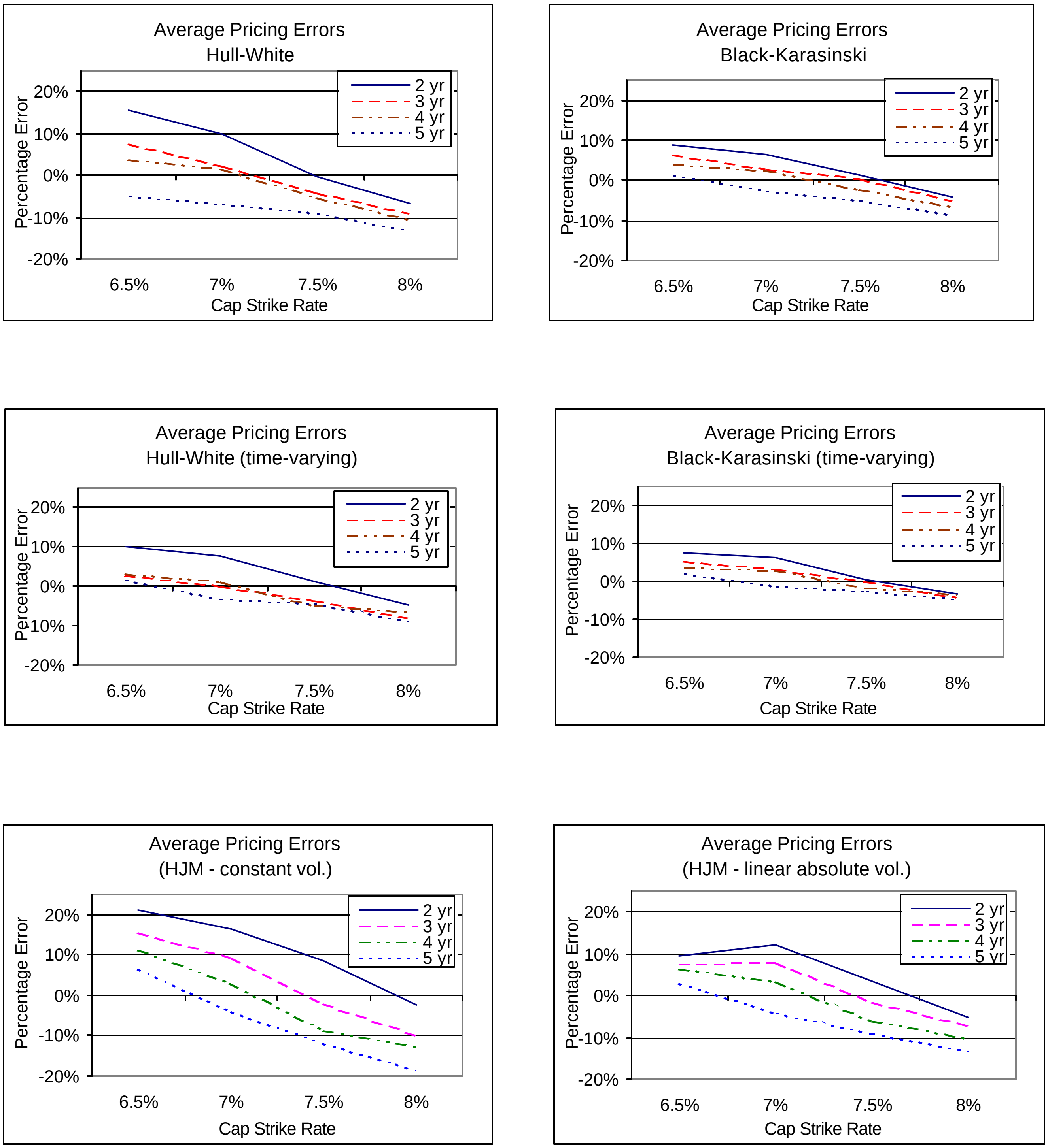
Table 8

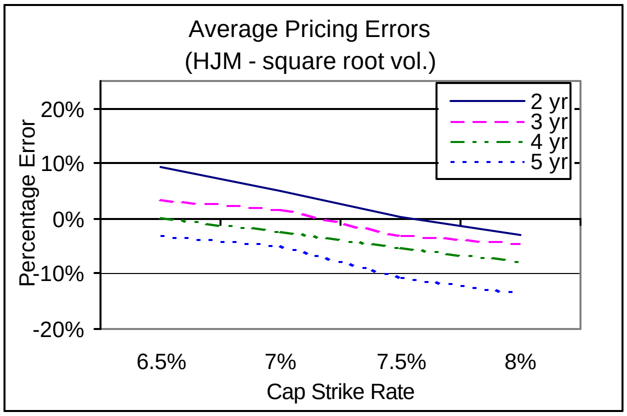
This table presents summary statistics for the hedging errors for the one-factor and two-factor spot rate and forward rate models analyzed. The hedging error is defined as the percentage change in the value of the hedge portfolio over a 5-day and a 20-day rebalancing interval. This error is averaged over the 219 days (March-December, 1998) for which the study was done. The hedge portfolio consists of one each of all the caps (floors) in the sample, across the four strike rates and the four maturities, and the appropriate Eurodollar futures contracts.

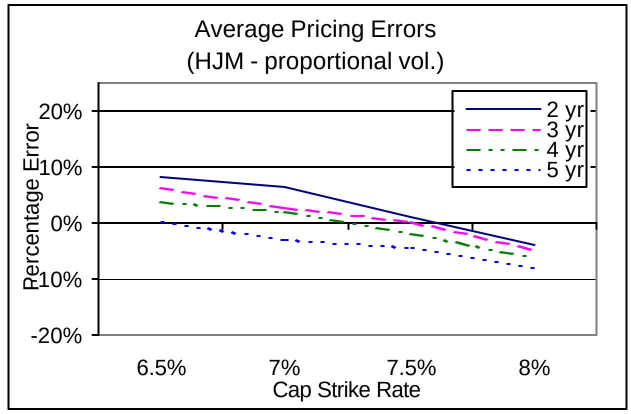
Model	Caps				Floors			
	5-day rebal.		20-day rebal.		5-day rebal.		20-day rebal.	
	Avg. % Error	Avg. % Abs Error	Avg. % Error	Avg. % Abs Error	Avg. % Error	Avg. % Abs Error	Avg. % Error	Avg. % Abs Error
Spot Rate Models								
Hull and White	0.05%	0.56%	0.17%	2.67%	0.03%	0.76%	0.22%	3.04%
HW - time varying	0.04%	0.51%	0.29%	3.22%	0.05%	0.59%	0.35%	4.22%
Black & Karasinski	-0.03%	0.41%	-0.09%	2.05%	0.06%	0.58%	0.19%	2.41%
BK - time varying	-0.12%	0.32%	0.03%	2.11%	0.11%	0.53%	0.25%	2.78%
Forward Rate Models – One Factor								
Absolute	0.08%	0.68%	0.07%	2.44%	0.12%	0.81%	0.04%	3.15%
Linear Absolute	0.11%	0.52%	0.13%	2.23%	0.09%	0.75%	0.14%	2.57%
Square Root	0.10%	0.46%	0.21%	1.98%	-0.13%	0.44%	-0.08%	2.16%
Proportional	0.04%	0.33%	0.07%	1.62%	0.07%	0.31%	0.11%	1.55%
Linear Proportional	0.04%	0.37%	0.08%	1.67%	0.05%	0.29%	0.09%	1.69%
Forward Rate Models – Two Factor								
Absolute	0.02%	0.19%	0.05%	0.54%	0.01%	0.16%	0.02%	0.74%
Proportional	0.02%	0.11%	0.04%	0.47%	-0.02%	0.15%	-0.01%	0.49%
Market Model - One Factor								
BGM	0.05%	0.38%	0.07%	1.65%	0.06%	0.33%	0.12%	1.59%

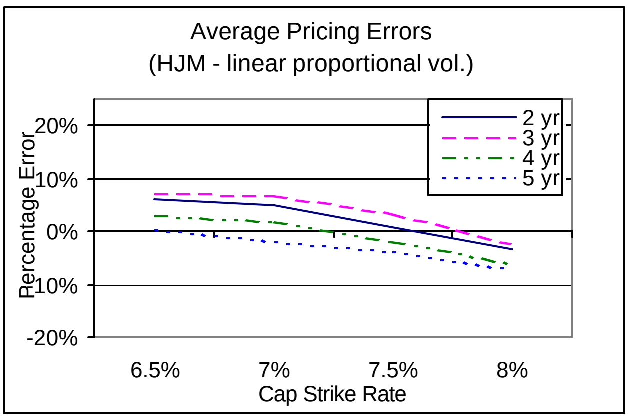
Figure 1

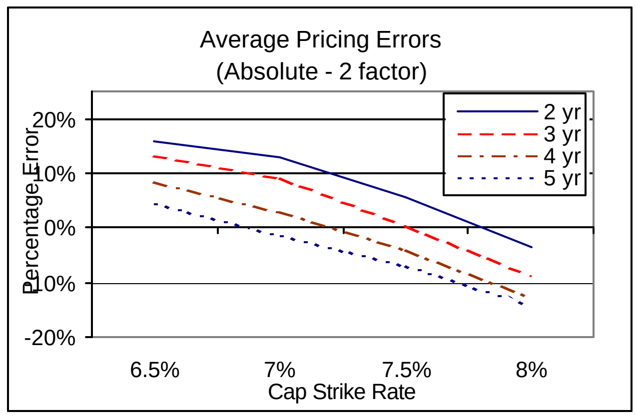
These figures present the average percentage pricing errors in predicting the prices of caps, using the one-factor and two-factor spot rate and forward rate models. The errors presented pertain to caps of 2-, 3-, 4- and 5-year maturity for strike rates of 6.5%, 7%, 7.5% and 8%. These errors are averaged over the 219 trading day sample period, March 1 - December 31, 1998.

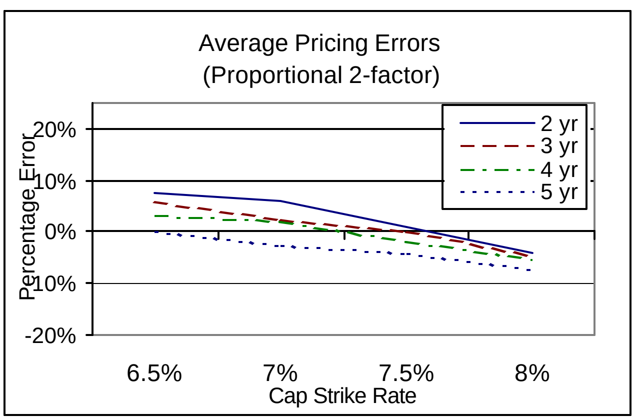



Average Pricing Errors
(HJM - square root vol.)
Percentage Error
20%
10%
0%
-10%
-20%
6.5%
7%
7.5%
8%
Cap Strike Rate
2 yr
3 yr
4 yr
5 yr


Average Pricing Errors
(HJM - proportional vol.)
Percentage Error
20%
10%
0%
-10%
-20%
6.5%
7%
7.5%
8%
Cap Strike Rate
2 yr
3 yr
4 yr
5 yr


Average Pricing Errors
(HJM - linear proportional vol.)
Percentage Error
20%
10%
0%
-10%
-20%
6.5%
7%
7.5%
8%
Cap Strike Rate
2 yr
3 yr
4 yr
5 yr


Average Pricing Errors
(Absolute - 2 factor)
Percentage Error
20%
10%
0%
-10%
-20%
6.5%
7%
7.5%
8%
Cap Strike Rate
2 yr
3 yr
4 yr
5 yr


Average Pricing Errors
(Proportional 2-factor)
Percentage Error
20%
10%
0%
-10%
-20%
6.5%
7%
7.5%
8%
Cap Strike Rate
2 yr
3 yr
4 yr
5 yr

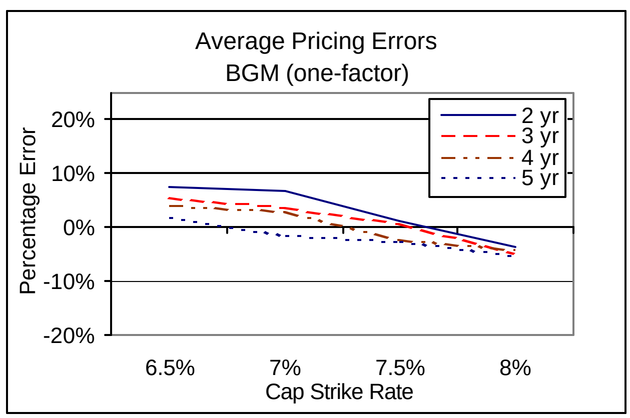
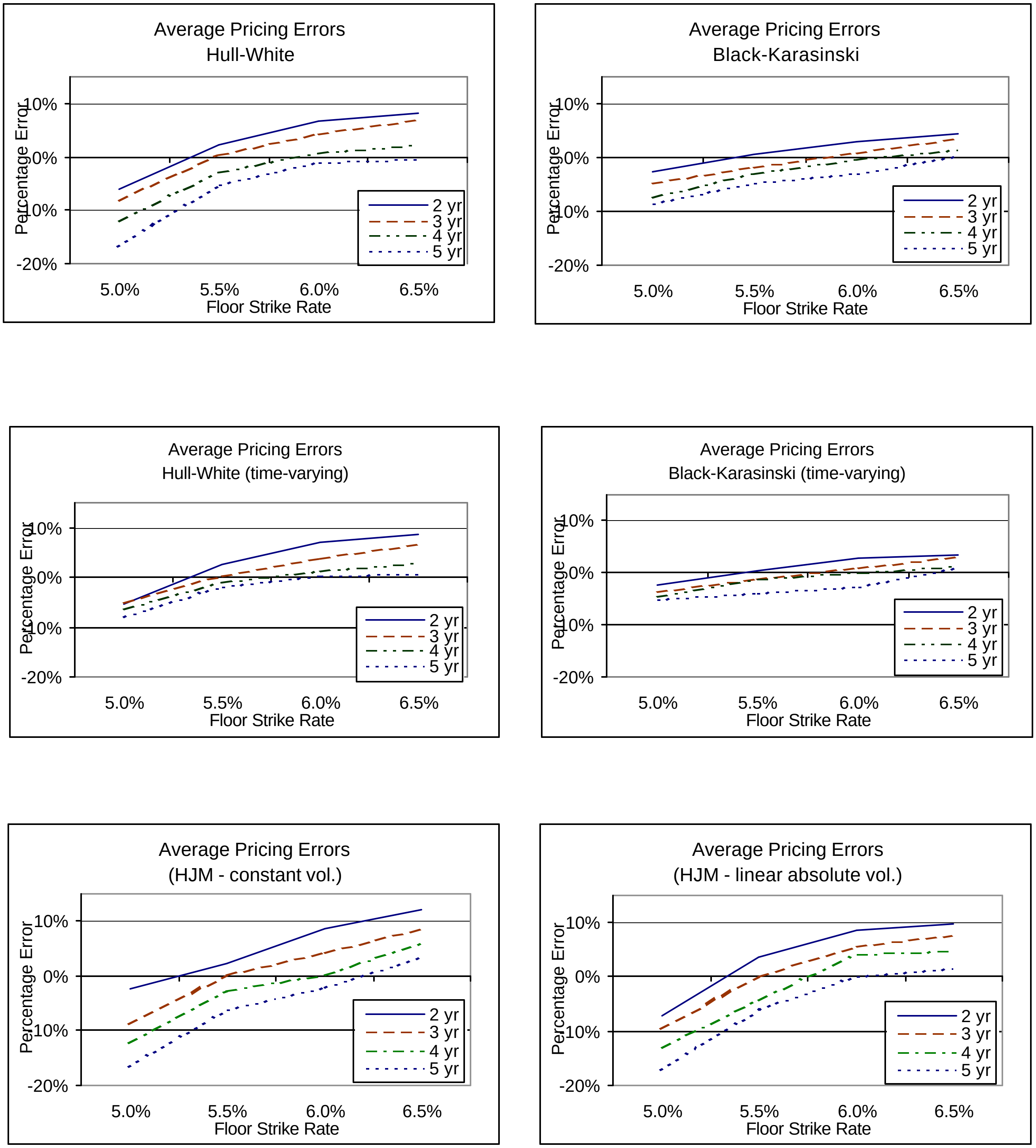
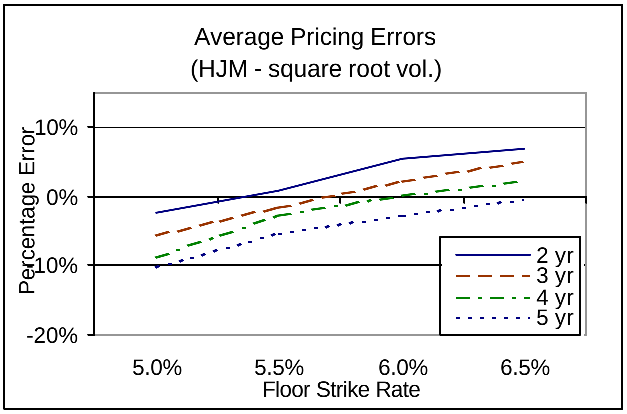
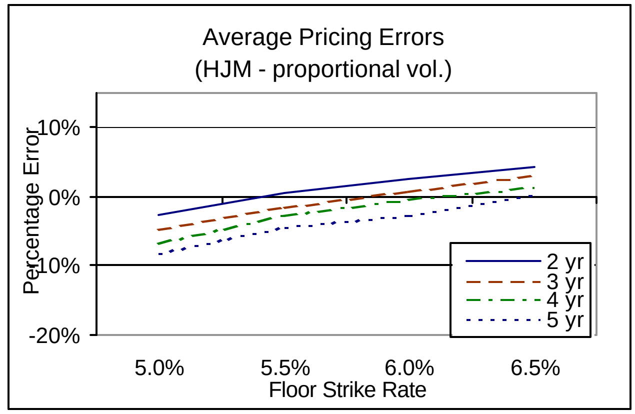

Average Pricing Errors
BGM (one-factor)
Percentage Error
20%
10%
0%
-10%
-20%
6.5%
7%
7.5%
8%
Cap Strike Rate
2 yr
3 yr
4 yr
5 yr

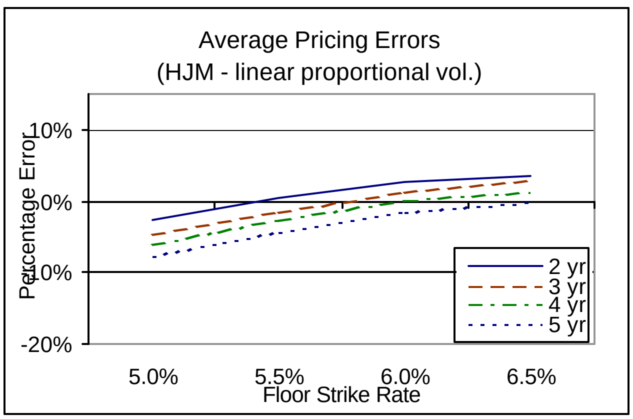
Figure 2

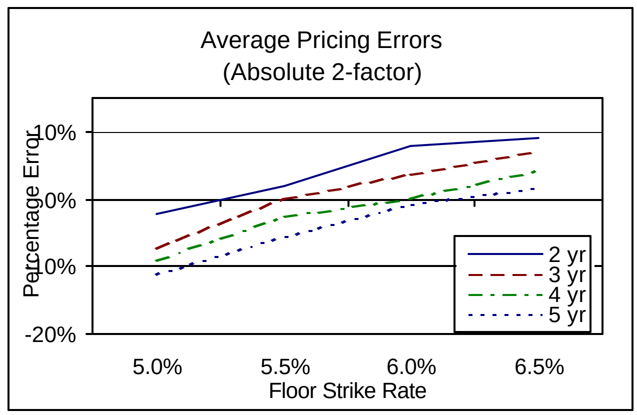
These figures present the average percentage pricing errors in predicting the prices of floors, using the spot rate and forward rate models. The errors presented pertain to floors of 2-, 3-, 4- and 5-year maturity for strike rates of 5%, 5.5%, 6% and 6.5%. These errors are averaged over the 219 trading day sample period, March 1 - December 31, 1998.

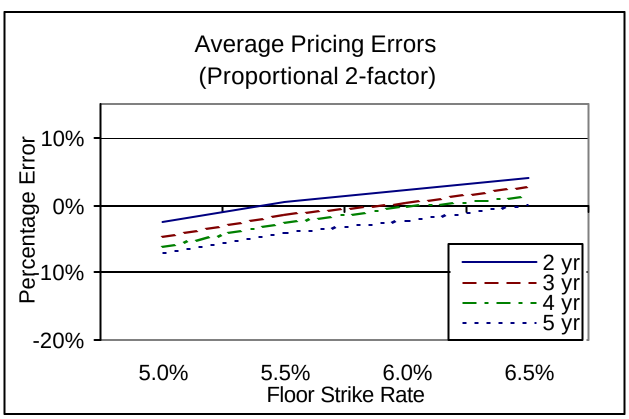


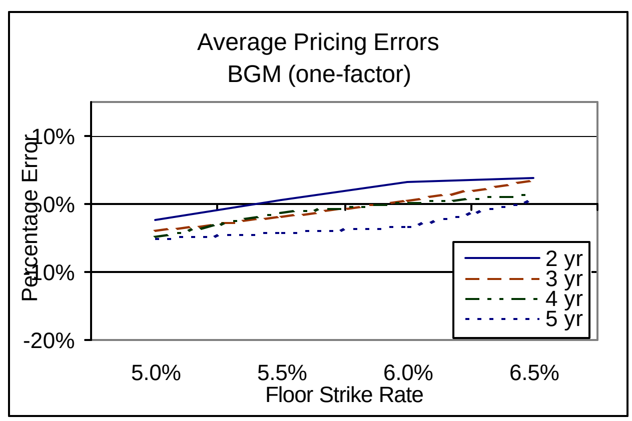

Average Pricing Errors
(HJM - square root vol.)
Percentage Error
10%
0%
-10%
-20%
5.0%
5.5%
6.0%
6.5%
Floor Strike Rate
2 yr
3 yr
4 yr
5 yr


Average Pricing Errors
(HJM - proportional vol.)
Percentage Error
10%
0%
-10%
-20%
5.0%
5.5%
6.0%
6.5%
Floor Strike Rate
2 yr
3 yr
4 yr
5 yr


Average Pricing Errors
(HJM - linear proportional vol.)
Percentage Error
10%
0%
-10%
-20%
5.0%
5.5%
6.0%
6.5%
Floor Strike Rate
2 yr
3 yr
4 yr
5 yr


Average Pricing Errors
(Absolute 2-factor)
Percentage Error
10%
0%
-10%
-20%
5.0%
5.5%
6.0%
6.5%
Floor Strike Rate
2 yr
3 yr
4 yr
5 yr


Average Pricing Errors
(Proportional 2-factor)
Percentage Error
10%
0%
-10%
-20%
5.0%
5.5%
6.0%
6.5%
Floor Strike Rate
2 yr
3 yr
4 yr
5 yr


Average Pricing Errors
BGM (one-factor)
Percentage Error
10%
0%
-10%
-20%
5.0%
5.5%
6.0%
6.5%
Floor Strike Rate
2 yr
3 yr
4 yr
5 yr